

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:23:43 ; Search time 26.01 Seconds
(Without alignments)
4330.246 Million cell updates/sec

Title: US-08-212-185-12

Sequence: 1 MAQWNOLOOLDRYLKQLHQ.....OPESLTFMDLSECATSPM 770

Scoring table: Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: SPREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_proteint:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	532	69.1	770	4 014916	014916 homo sapien
2	473	61.4	769	4 09BM54	09BM54 homo sapien
3	271	35.2	769	13 09PVX8	09PVX8 xenopus lae
4	145	18.8	163	6 09N145	09N145 macaca mula
5	63	8.2	394	13 09DDJ8	09DDJ8 brachydanio
6	63	8.2	806	13 093599	093599 brachydanio
7	59	7.7	767	13 013133	013133 oncorhynch
8	56	7.3	56	6 09N034	09N034 bos taurus
9	51	6.6	51	11 099ML3	099ML3 mus musculu
10	43	5.6	43	11 020XR4	020XR4 rattus sp.
11	38	4.9	51	4 09BMX2	09BMX2 homo sapien
12	19	2.5	754	13 013132	013132 oncorhynch
13	12	1.6	43	11 09QVR3	09QVR3 rattus sp.
14	12	1.6	156	4 09UDL5	09UDL5 homo sapien
15	12	1.6	712	11 099K94	099K94 mus musculu
16	12	1.6	749	11 09QXK0	09QXK0 rattus norv
17	12	1.6	749	11 09D323	09D323 mus musculu
18	12	1.6	754	13 013131	013131 oncorhynch
19	11	1.4	155	4 09UDL4	09UDL4 homo sapien

20	11	1.4	195	6 09GKX9	09GKX9 sus scrofa
21	11	1.4	749	13 093598	093598 brachydanio
22	9	1.2	44	5 09XXE3	09XXE3 caenorhabdi
23	9	1.2	72	5 018168	018168 caenorhabdi
24	9	1.2	732	5 09VY71	09VY71 drosophila
25	9	1.2	922	11 09QXJ2	09QXJ2 mus musculu
26	9	1.2	925	11 09Q2E4	09Q2E4 mus musculu
27	8	1.0	85	6 029340	029340 sus scrofa
28	8	1.0	111	11 070406	070406 rattus norv
29	8	1.0	141	11 070429	070429 rattus norv
30	8	1.0	177	6 09N0E5	09N0E5 bos taurus
31	8	1.0	293	2 09KRP8	09KRP8 vibrio chol
32	8	1.0	348	11 099K06	099K06 mus musculu
33	8	1.0	515	10 09FKI8	09FKI8 arabidopsis
34	8	1.0	581	2 087594	087594 streptomyce
35	8	1.0	588	2 09RJ58	09RJ58 streptomyce
36	8	1.0	592	13 09PWP7	09PWP7 gallus gall
37	8	1.0	679	4 09BQD2	09BQD2 homo sapien
38	8	1.0	722	5 097164	097164 anopheles g
39	8	1.0	728	1 09HOB3	09HOB3 halobacteri
40	8	1.0	784	13 09DC74	09DC74 oncorhynch
41	8	1.0	786	11 09JXM1	09JXM1 mus musculu
42	8	1.0	787	13 093378	093378 gallus gall
43	8	1.0	843	12 088944	088944 kapoti's sa
44	8	1.0	1088	12 055590	055590 avian rotav
45	8	1.0	1088	12 085036	085036 porcine rot

ALIGNMENTS

RESULT 1
ID 014916 PRELIMINARY; PRT; 770 AA.
AC 014916:
DT 01-JAN-1998 (TREMBLrel, 05, Created)
DT 01-AUG-1999 (TREMBLrel, 11, Last sequence update)
DT 01-JUN-2001 (TREMBLrel, 17, Last annotation update)
DE TRANSCRIPTION FACTOR (SIGNAL, TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION
3).
GN STAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98296260; PubMed=9630560;
RA Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT "Highly conserved amino-acid sequence between murine STAT3 and a
revised human STAT3 sequence.";
RL Gene 213:119-124(1998).
RN [2]
RP SEQUENCE OF 564-704 FROM N.A.
RC TISSUE=HEPATOMA;
RA Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL Submitted (Oct-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ012463; CAAL0032.1; -;
DR EMBL; AF029311; AAB84254.1; -;
DR HSSP; P42227; IBG1.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 2.
SQ SEQUENCE 770 AA; 88067 MM; 6C00632211C8012D CRC64;

Query Match 69.1%; Score 532; DB 4; Length 770;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

[illegible]

Query Match	61.4%	Score 473	DB 4	Length 769	
Best Local Similarity	99.7%	Pred. No. 0			
Matches 673	Conservative	0	Mismatches 2	Indels 0	Gaps 0
QY 26	PFMELROFLAMISODMAYAAKSESHATLVFHNLLGELIDDOYSRFLIDESNLYOHNLRR	85			
Db 26	PFMELROFLAMIESODMAYAAKSESHATLVFHNLLGELIDDOYSRFLIDESNLYOHNLRR	85			
QY 86	IKOFLQSHYLEKPEMEIARIIVAKCIWEESSHLIOTATAOOGQAMHPPIAAVTEKQOMLE	145			
Db 86	IKOFLQSHYLEKPEMEIARIIVAKCIWEESSRLIOTATAOOGQAMHPPIAAVTEKQOMLE	145			
QY 146	OHLDVRRKVOYDEQKMKVVENLODDPFPNKTKLSQSGMODLKNNGSVTRQKKMOLEQ	205			
Db 146	OHLDVRRKVOYDEQKMKVVENLODDPFPNKTKLSQSGMODLKNNGSVTRQKKMOLEQ	205			
QY 206	MLTALDQMRSTIVSELGALLSMEYVOKTITDEELADMKRPEIACIGGPNICIDRLLEN	265			
Db 206	MLTALDQMRSTIVSELGALLSMEYVOKTITDEELADMKRROUICIGGPNICIDRLLEN	265			
QY 266	WITSIAESQLOTRQOIKKLELEQOVSYKGDPIYOHRIPLERIVELFENLKKSAFYVER	325			
Db 266	WITSIAESQLOTRQOIKKLELEQOVSTKGDPIYOHRIPLERIVELFENLKKSAFYVER	325			
QY 326	QPCMPMHPDRPLVITKGVQFTTKVRLVLPPELNTYOLKIKVCIDKDSGDVAALRGSKKFN	385			
Db 326	QPCMPMHPDRPLVITKGVQFTTKVRLVLPPELNTYOLKIKVCIDKDSGDVAALKSKKFN	385			
QY 386	ILGNTTKVMNNEESNGSLSAFFKRLTIREQCGNGANGANDASLIIVTEHLITFEFEV	445			
Db 386	ILGNTTKVMNNEESNGSLSAFFKRLTIREQCGNGANGANDASLIIVTEHLITFEFEV	445			
QY 446	YHOGKIDLETHSLPVVYISNICQMPNMAWILWYNMLTNNKKNVNFPTKPIGTMDOVA	505			
Db 446	YHOGKIDLETHSLPVVYISNICQMPNMAWILWYNMLTNNKKNVNFPTKPIGTMDOVA	505			
QY 506	EVLWSQSFSTTKRGSLTQLTTLAEKLLGPGVNYSGCOITMAKPFCKENNAGGFSFWWL	565			
Db 506	EVLWSQSFSTTKRGSLTQLTTLAEKLLGPGVNYSGCOITMAKPFCKENNAGGFSFWWL	565			
QY 566	DNIIIDLVRKYILALNNEGIMGFIKSERERALLSTRPGCTFLRLRESSEKEGGVFTWE	625			
Db 566	DNIIIDLVRKYILALNNEGIMGFIKSERERALLSTRPGCTFLRLRESSEKEGGVFTWE	625			
QY 626	KDISKTOIQSEVPTTKOOLNNNSFAETIMGKINDATILVSPILVYLPDIPKEKFAEQ	685			
Db 626	KDISKTOIQSEVPTTKOOLNNNSFAETIMGKINDATILVSPILVYLPDIPKEKFAEQ	685			
QY 686	YCRPESQEHPEADPG 700				
Db 686	YCRPESQEHPEADPG 700				
RESULT 3					
Q9PVX8	PRELIMINARY:	PRT: 769 AA.			
AC Q9PVX8:					
DT 01-MAY-2000 (Tremblrel. 13, Created)					
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)					
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)					
DE STAT 3.					
GN STAT 3.					
OS Xenopus laevis (African clawed frog).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;					
OC Xenopodinae; Xenopus.					
OX NCBI_TaxID=8335;					
RM [1]					
RP SEQUENCE FROM N.A. Matsunoto Y., Matsuda T., Arizumi T., Heike T.,					
RA Nishinakamura R., Asashima M., Yokota T.;					
*Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus					

DR HSSP: P42227; 1BG1.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS0001; SH2; 1.
 SO SEQUENCE 806 AA; 92151 MW; 7ABCAEAA01C3C942 CRC64;

Query Match 8.2%; Score 63; DB 13; Length 806;
 Best Local Similarity 100.0%; Pred. No. 1.5e-57;
 Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 537 VNYSQGITMAKFKENNAKGFSFWVLDNITDLYKRYIALALNMGYIMGFISKERERA 596
 DB 538 VNYSQGITMAKFKENNAKGFSFWVLDNITDLYKRYIALALNMGYIMGFISKERERA 597
 OY 597 ILS 599
 DB 598 ILS 600

RESULT 7
 O13133 PRELIMINARY: PRT: 767 AA.
 AC O13133:
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE STAT3.
 GN RBTSTAT3.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Procaantopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Johnson M.C., Mourich D.V., Leong J.C.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U60333; AAB60926.1; -.
 DR HSSP: P42227; 1BG1.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 SO SEQUENCE 767 AA; 87816 MW; FED97740C74C3798 CRC64;

Query Match 7.7%; Score 59; DB 13; Length 767;
 Best Local Similarity 100.0%; Pred. No. 2.5e-53;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 427 DASLIVTEELHITFETEVYHOGKIDLETHSLPVVVISNICOHPNANASILMYMLTN 485
 DB 428 DASLIVTEELHITFETEVYHOGKIDLETHSLPVVVISNICOHPNANASILMYMLTN 486

RESULT 8
 O9N0J4 PRELIMINARY: PRT: 56 AA.
 AC O9N0J4:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3, STAT3 (FRAGMENT).
 GN STAT3.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Molenaar A., Wheeler T.T., McCracken J.Y., Seyfert H.M.;
 RT "The STAT3-encoding gene resides within the 40 kbp gap between the
 RT STAT5A- and STAT5B-encoding genes in cattle."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ276489; CAB93140.1; -.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF01017; STAT; 1.
 FT NON_TER 1
 FT NON_TER 56
 SO SEQUENCE 56 AA; 6057 MW; 4450AC5DD0BD64A CRC64;

Query Match 7.3%; Score 56; DB 6; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3.6e-51;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 372 SGDVALMGSKRKNITIGNTKVMNNEESNNGSLAEFKHLTLREORCGNGGRAND 427
 DB 1 SGDVALMGSKRKNITIGNTKVMNNEESNNGSLAEFKHLTLREORCGNGGRAND 56

RESULT 9
 O99ML3 PRELIMINARY: PRT: 51 AA.
 AC O99ML3:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE STAT3 (FRAGMENT).
 GN STAT3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shao H., Tweardy D.J.;
 RT "Murine Stat3 gene exon 22 to exon 23."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF332507; AAK17195.1; -.
 DR NON_TER 1
 FT NON_TER 51
 SO SEQUENCE 51 AA; 5379 MW; CE2B44AD932D12 CRC64;

Query Match 6.6%; Score 51; DB 13; Length 51;
 Best Local Similarity 100.0%; Pred. No. 6.9e-46;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 702 AAPYLKTFKFCVPTTGCSNFTIDLPMSPTLDLSLQFGNNGEGAEPSAGGF 752
 DB 1 AAPYLKTFKFCVPTTGCSNFTIDLPMSPTLDLSLQFGNNGEGAEPSAGGF 51

RESULT 10
 O9QVR4 PRELIMINARY: PRT: 43 AA.
 AC O9QVR4:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE INTERLEUKIN-6 RESPONSE ELEMENT BINDING PROTEIN (FRAGMENT).
 OS Rattus sp.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95398099; PubMed=7545365;
 RA Ripberger J., Fritz S., Richter K., Dreier B., Schneider K.,
 Lochner K., Marschalek R., Hocke G., Lottspeich F., Fey G.H.;


```
RT      *isolation of two interleukin-6 response element binding proteins from
RT      acute phase rat livers.*
RL      Ann. N. Y. Acad. Sci. 762:252-260(1995).
DR      HSSP: P42227: 1BG1.
DR      InterPro: IPR000980: SH2.
DR      Pfam: PF00017: SH2: 1.
DR      PROSITE: PS50001: SH2: 1.
DR      PROSITE: PS50001: SH2: 1.
SQ      SEQUENCE 43 AA: 4868 MW: 0C2F14721B863798 CRC64:

Query Match
Best Local Similarity 100.0%; Score 43; DB 11; Length 43;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY      581 NEGYIMGFISKERERAILSTKPGCTFLLRFSSESKEGCVFTW 623
Db      1 NEGYIMGFISKERERAILSTKPGCTFLLRFSSESKEGCVFTW 43

RESULT 11
O9BXH2 ID O9BXH2 PRELIMINARY: PRT: 51 AA.
AC O9BXH2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE INTERLEUKIN-6, RESPONSE ELEMENT BINDING PROTEIN (FRAGMENT).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shao H., Tweardy D.J.:
RT "Human STAT3 gene exon 22 to exon 23."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF332508; AAK17196.1; -.
FT NON_TER 1 51
FT NON_TER 1 51
SQ SEQUENCE 51 AA: 5375 MW: CE2B444D8AFC3D12 CRC64:

Query Match
Best Local Similarity 100.0%; Score 38; DB 4; Length 51;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY      702 AAPYLKTKFICVPTGCSNTIDLPMSRPTLDSLMQFCN 739
Db      1 AAPYLKTKFICVPTGCSNTIDLPMSRPTLDSLMQFCN 38

RESULT 12
O13132 ID O13132 PRELIMINARY: PRT: 754 AA.
AC O13132;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE STAT1-2.
DE STAT1-2.
GN RARSTAT1-2.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Proacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnson M.C., Mouritch D.V., Leong J.C.:
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U60332; AAB60925.1; -.
DR HSSP: P42224: 1BF5.
RT InterPro: IPR000980: SH2.

DR      InterPro: IPR001217: STAT.
DR      Pfam: PF00017: SH2: 1.
DR      Pfam: PF01017: STAT: 1.
DR      SMART: SM00252: SH2: 1.
DR      PROSITE: PS50001: SH2: 1.
DR      PROSITE: PS50001: SH2: 1.
SQ      SEQUENCE 754 AA: 87147 MW: 340B9645EA040142 CRC64:

Query Match
Best Local Similarity 100.0%; Score 19; DB 13; Length 754;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY      382 RKNILCTNNTKVMNMEESN 400
Db      373 RKNILCTNNTKVMNMEESN 391

RESULT 13
O9QVR3 ID O9QVR3 PRELIMINARY: PRT: 43 AA.
AC O9QVR3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE INTERLEUKIN-6, RESPONSE ELEMENT BINDING PROTEIN (FRAGMENT).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RA Ripberger J., Fritz S., Richter K., Dreier B., Schneider K.,
RA Lochner K., Marschalek R., Hocke G., Lotzpeich F., Fey G.H.:
RT "Isolation of two interleukin-6 response element binding proteins from
RT acute phase rat livers."
RL Ann. N. Y. Acad. Sci. 762:252-260(1995).
DR HSSP: P42224: 1BF5.
DR InterPro: IPR000980: SH2.
DR Pfam: PF00017: SH2: 1.
DR PROSITE: PS50001: SH2: 1.
SQ SEQUENCE 43 AA: 4937 MW: 5D710A6749125C3B CRC64:

Query Match
Best Local Similarity 100.0%; Score 12; DB 11; Length 43;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

OY      585 IMGFISKEKERA 596
Db      5 IMGFISKEKERA 16

RESULT 14
O9UDL5 ID O9UDL5 PRELIMINARY: PRT: 156 AA.
AC O9UDL5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE INTERFERON-STIMULATED GENE FACTOR 3 ALPHA 91/84 KDA PROTEIN
DE (FRAGMENTS).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92346719; PubMed=1638633;
RA Fu X.Y.:
RT "A transcription factor with SH2 and SH3 domains is directly activated
RT by an interferon alpha-induced cytoplasmic protein tyrosine
RT kinase(s)."
```

RL Cell 70:323-335(1992).
 DR HSSP: P42224; 1BF5.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 FT NON_TER 1
 FT NON_CONS 54
 FT NON_TER 156
 SQ SEQUENCE 156 AA; 18004 MW; 4D6EAD7D353C0AAA CRC64;

Query Match 1.6%; Score 12; DB 4; Length 156;
 Best Local Similarity 100.0%; Pred. No. 0.00051;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 585 IMGFISKERERA 596
 |||||
 DB 60 IMGFISKERERA 71

RESULT 15
 O99K94 PRELIMINARY: PRT: 712 AA.
 AC O99K94;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE SIMILAR TO SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC004808; AAH04808.1; -;
 SQ SEQUENCE 712 AA; 83106 MW; D62081709638CDA5 CRC64;

Query Match 1.6%; Score 12; DB 11; Length 712;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 585 IMGFISKERERA 596
 |||||
 DB 578 IMGFISKERERA 589

Search completed: March 19, 2002, 15:27:46
 Job time: 243 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:22:48 ; Search time 14.48 Seconds

(Without alignments)
1949.718 Million cell updates/sec

Title: US-08-212-185-12

Sequence: 1 MAQWNLQGLDRLRYLKQLHQ.....QFESLTFMDLITSECATSPM 770

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	543	70.5	770	STA3_MOUSE	P42227 mus musculu
2	459	59.6	770	STA3_RAT	P52631 rattus norv
3	221	28.7	770	STA3_HUMAN	P40263 homo-sapien
4	12	1.6	749	STA1_MOUSE	P42225 mus musculu
5	12	1.6	750	STA1_HUMAN	P42224 homo sapien
6	11	1.4	748	STA4_HUMAN	Q14765 homo sapien
7	11	1.4	749	STA4_MOUSE	P42228 mus musculu
8	11	1.4	851	STA2_HUMAN	P52630 homo sapien
9	9	1.2	864	STA2_PIG	Q02799 sus scrofa
10	9	1.2	923	STA2_MOUSE	Q9W12 mus musculu
11	8	1.0	458	HDNO_ARTOX	P08159 arthobacte
12	8	1.0	663	MNEL_YEAST	P24720 saccharomyc
13	8	1.0	761	STAT_DROME	Q24151 drosophila
14	8	1.0	766	STA5_HUMAN	P51692 homo sapien
15	8	1.0	766	STA5_MOUSE	P42232 mus musculu
16	8	1.0	786	STA5_RAT	P52632 rattus norv
17	8	1.0	787	STA5_BOVIN	Q9LUM3 bos taurus
18	8	1.0	787	STA5_PIG	Q9LUM3 sus scrofa
19	8	1.0	793	STA5_MOUSE	P42230 mus musculu
20	8	1.0	793	STA5_RAT	Q62771 rattus norv
21	8	1.0	794	STA5_BOVIN	Q95115 bos taurus
22	8	1.0	794	STA5_HUMAN	P42229 homo sapien
23	8	1.0	794	STA5_SHEEP	P42231 ovis aries
24	8	1.0	799	STA5_PIG	Q9LUM1 sus scrofa
25	8	1.0	837	STA6_MOUSE	P52633 mus musculu
26	8	1.0	847	STA6_HUMAN	P42226 homo sapien
27	8	1.0	1088	RRPO_ROTBR	P17468 bovine rota
28	8	1.0	1088	RRPO_ROTBU	P21615 bovine rota
29	8	1.0	1088	RRPO_ROTGS	P17699 porcine rota
30	8	1.0	1088	RRPO_ROTSL	P22678 simian 11 r
31	7	0.9	74	YD77_MYCPN	Q9EX03 mycoplasma
32	7	0.9	75	RL29_MYCBO	Q06030 mycobacteri
33	7	0.9	77	RL29_MYCTU	P95057 mycobacteri

ALIGNMENTS

RESULT	ID	STANDARD	PRT	770 AA	
1	STA3_MOUSE				
AC	P42227				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3 (ACUTE-PHASE RESPONSE FACTOR).				
GN	STAT3 OR APRF.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 154-158; 181-185 AND 632-640.				
RC	STRAIN=BALB/C; TISSUE=Liver;				
RX	MEDLINE=94208062; PubMed=7512451;				
RA	Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,				
RA	Yoshida K., Sudo T., Naruto M., Kishimoto T.,				
RT	"Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-				
RT	related transcription factor involved in the gp130-mediated signaling				
RT	pathway."				
RL	Cell 77:63-71(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Thymus;				
RX	MEDLINE=94188718; PubMed=8140422;				
RA	Zhong Z., Wen Z., Darnell J.E., Jr.;				
RT	"Stat3: a STAT family member activated by tyrosine phosphorylation in				
RT	response to epidermal growth factor and interleukin-6."				
RL	Science 264:95-98(1994).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RX	MEDLINE=95014185; PubMed=7523373;				
RA	Raz R., Durbin J.E., Levy D.E.;				
RT	"Acute phase response factor and additional members of the				
RT	interferon-stimulated gene factor 3 family integrate diverse signals				
RT	from cytokines, interferons, and growth factors."				
RL	J. Biol. Chem. 269:24391-24395(1994).				
RN	[4]				
RP	SEQUENCE FROM N.A. (STAT3B).				
RC	STRAIN=BALB/C; AND B6; TISSUE=Liver;				
RX	MEDLINE=96016116; PubMed=7568080;				
RA	Schaeter T.S., Sanders L.K., Nathans D.;				
RT	"Cooperative transcriptional activity of Jun and Stat3 beta, a short				
RT	form of Stat3."				
RL	Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).				
RN	[5]				
RP	PHOSPHORYLATION OF SER-727, AND MUTAGENESIS.				
RX	MEDLINE=95354205; PubMed=7543024;				
RA	Wen Z., Zhong Z., Darnell J.E., Jr.;				
RT	"Maximal activation of transcription by Stat1 and Stat3 requires both				
RT	tyrosine and serine phosphorylation."				
RL	Cell 82:241-250(1995).				

34	7	0.9	94	1	SPDA_STRLI	P22407 streptomyc
35	7	0.9	128	1	VF70_AQUAE	O67514 aquilex aeo
36	7	0.9	136	1	SR14_HUMAN	P37108 homo sapien
37	7	0.9	161	1	CC31_YEAST	P06704 saccharomyc
38	7	0.9	170	1	Y586_METJA	O58006 methanococ
39	7	0.9	173	1	PSAL_MASJA	O11126 mastigoclad
40	7	0.9	193	1	ES31_DROME	P48593 drosophilad
41	7	0.9	205	1	MD21_HUMAN	Q13257 homo sapien
42	7	0.9	205	1	MD21_MOUSE	Q92105 mus musculu
43	7	0.9	209	1	YR8K_HABIN	P45075 haemophilus
44	7	0.9	221	1	ERG2_MACGR	P33281 magnaporthe
45	7	0.9	238	1	AGRA_STAAU	P13131 staphylococ

RN	{6}
RP	X-RAY CRYSTALLOGRAPHY (2 25 ANGSTROMS) OF 136-716.
RX	MEDLINE=98334373; PubMed=9671298;
RA	Becker S., Groner B., Mueller C.W.:
RT	"Three-dimensional structure of the Stat3beta homodimer bound to DNA.";
RL	Nature 394:145-151(1998).
CC	-I- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6 (IL-6)-RESPONSIVE ELEMENT IDENTIFIED IN THE PROMOTERS OF VARIOUS ACUTE-PHASE PROTEIN GENES. STAT3B INTERACTS WITH THE N-TERMINAL PART OF C-JUN TO ACTIVATE SUCH PROMOTERS IN A COOPERATIVE WAY.
CC	-I- PATHWAY: INVOLVED IN THE GPI30-MEDIATED SIGNALING PATHWAY.
CC	-I- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY MEMBER (AT LEAST STAT1) (BY SIMILARITY).
CC	-I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN RESPONSE TO PHOSPHORYLATION.
CC	-I- ALTERNATIVE PRODUCTS: 2 ISOFORMS: STAT3A (SHOWN HERE) AND STAT3B; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	-I- TISSUE SPECIFICITY: STAT3A IS SEEN IN THE LIVER, SPLEEN, AND KIDNEY. STAT3B IS ALSO DETECTED IN THE LIVER, ALTHOUGH IN A MUCH LESS ABUNDANT MANNER.
CC	-I- PPM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNF, LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3 HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC	-I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC	-I- SIMILARITY: CONSTAINS 1 SH2 DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isdb.ch/announce/ or send an email to license@isdb.stb.ch).
CC	-----
DR	EMBL: L29278; AAA37254.1; -
DR	EMBL: U06922; AAA19452.1; -
DR	EMBL: U08378; AAA56668.1; -
DR	EMBL: U30709; AAC52612.1; -
DR	PDB: 1BG1; 1J-JAN-99.
DR	TRANSFAC: T01574; -
DR	MCD: MG1:103038; Stat3.
DR	InterPro: IPR000980; SH2.
DR	Interpro: IPR001217; STAT.
DR	pfam: PF00017; SH2_1.
DR	SMART: PF01017; STAT_1.
DR	SMART: SM00252; SH2_1.
DR	PROSITE: PS00001; SH2; 1.
KW	Transcription regulation; DNA-binding; Nuclear protein; SH2 domain; Phosphorylation; Alternative splicing; 3D-structure.
FT	DOMAIN 580 .. 670
FT	MOD_RES 705
FT	MOD_RES 705
FT	MOD_RES 705
FT	MOD_RES 727
FT	VARSPLIC 716 .. 770
FT	VARSPLIC 716
FT	FT
FT	VARIAINT 701 .. 701
FT	MUTAGEN 727 .. 727
FT	CONFLICT 16 .. 16
FT	CONFLICT 25 .. 25
FT	CONFLICT 394 .. 394
SO	SEQUENCE 770 AA; 88053 MW; 6C00626711C8012D CRC64;
Query Match	70.5%; Score 543; DB 1; Length 770;
Best local Similarity	99.7%;
Matches 743;	Conservative 0; Mismatches 2; Indels 0; Gaps 0;
oy	26 FFMELROFLAWISEQDNVAAKSKSHATLVFNHLIGFDIOOYSRFLOEINVLYOHNLRR 85

Db	26	FMELRQFLAPMIESQDMAVYAAASKSHATLVFHNILGCELDQY5KPLDJSVLYIQHNLJK	85
Qy	86	IKOFLQSRITLKPMEIARIIVARCLMEESRLLQITATAAOOGGQANHPIAAVYTEKQOMLE	1455
Db	86	IKOFLQSRITLKPMEIARIIVARCLMEESRLLQITATAAOOGGQANHPIAAVYTEKQOMLE	1455
Qy	146	OHLDQVRKRVODLPQKMKVVENLDODDFPNYKTLKSGQDMODLGNNO5YTRQKMOOLEQ	2050
Db	146	OHLDQVRKRVODLEBKMKVVENLDODDFPNYKTLKSGQDMODLGNNO5YTRQKMOOLEQ	2050
Qy	206	MILTADQMRRSIVSELAGLLSAMEYVOKTLTDEELADKKRPEIACIGCPNICTDRLFN	2655
Db	206	MILTADQMRRSIVSELAGLLSAMEYVOKTLTDEELADKKRPEIACIGCPNICTDRLFN	2655
Qy	266	WTTLSAESOLOTRQOIKKLEELQKVSXKSPRIVOIHRPMLERIVELFRNLKMSAFYVER	3255
Db	266	WTTLSAESOLOTRQOIKKLEELQKVSXKSPRIVOIHRPMLERIVELFRNLKMSAFYVER	3255
Qy	326	QPCMPMHDPRLVIKTVGOVFTTKMYLIVKPELNLVQIKIVKCIJDKDSQDVAALGSKRFN	3855
Db	326	QPCMPMHDPRLVIKTVGOVFTTKMYLIVKPELNLVQIKIVKCIJDKDSQDVAALGSKRFN	3855
Qy	386	ILGTTTKYMNNEESNGSLSAEPKHLITREORCGNGRANCASLIYTBELHLITFETEV	4455
Db	386	ILGTTTKYMNNEESNGSLSAEPKHLITREORCGNGRANCASLIYTBELHLITFETEV	4455
Qy	446	YHOGKLTLETHSLDVPVYVVISNIGCMPNMA5ILWYNMLTNPKNVNFFTRPIGTMQOVA	5050
Db	446	YHOGKLTLETHSLDVPVYVVISNIGCMPNMA5ILWYNMLTNPKNVNFFTRPIGTMQOVA	5050
Qy	506	EYLSQWFSSTTKRGCLIEQLTTLAEKLLGPGVNTSGCOITWAKFCCKENMAGKGF5VWML	5655
Db	506	EYLSQWFSSTTKRGCLIEQLTTLAEKLLGPGVNTSGCOITWAKFCCKENMAGKGF5VWML	5655
Qy	566	DNITDLYKKYILALMNEGYMGCFISKEBERA1LSTKPGTFLFFSESSKEGQVTFHWVE	6255
Db	566	DNITDLYKKYILALMNEGYMGCFISKEBERA1LSTKPGTFLFFSESSKEGQVTFHWVE	6255
Qy	626	KDISKTOIQOSEVPTKQOLNNMSPAEIIIMGYKIMDATNIIIVSPLVYLPDIPKEAEFGK	6855
Db	626	KDISKTOIQOSEVPTKQOLNNMSPAEIIIMGYKIMDATNIIIVSPLVYLPDIPKEAEFGK	6855
Qy	686	YCRPESQHPENDPESAPARYLKTFCICVTPPTG5NTIDLPMSPTDLSLQWQNGNGCAE	7455
Db	686	YCRPESQHPENDPESAPARYLKTFCICVTPPTG5NTIDLPMSPTDLSLQWQNGNGCAE	7455
Qy	746	PSAGQFESLTFDMDJTSECATSPM 770	
Db	746	PSAGQFESLTFDMDJTSECATSPM 770	
RESULT 2			
ID	STA3_RAT	STANDARD:	PRT: 770 AA.
AC	P52631:		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3.		
GN	STAT3.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxId=10116;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RX	MEDLINE=96102059; Pubmed=8530402;		
RA	Rippeger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,		
RA	Fey G.H.;		
RT	"Transcription factors Stat3 and Stat5b are present in rat liver		
RT	nuclei late in an acute phase response and bind interleukin-6		
RT	response elements";		

CC		This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (see http://www.isb.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC		-
DR	EMBL:	X91810; CAA62920.1; .
DR	HSSP:	P42224; 1hp5.
DR	InterPro:	IPR000980; SH2.
DR	InterPro:	IPR001217; STRAT.
DR	Pfam:	PF00017; SH2; 1.
DR	Pfam:	PF01017; STAT; 1.
DR	SMART:	SMO0252; SH2; 1.
DR	PROSITE:	PS50001; SH2; 1.
KW		Transcription regulation; DNA-binding; Nuclear protein;
RK	Phosphorylation:	SH2 domain..
FT	DOMAIN	580 670
FT	MOD_RES	705 705
FT		SH2.
FT		PHOSPHORYLATION (BY JAKS) (BY
FT		SIMILARITY).
FT	MOD_RES	727 727
FT		PHOSPHORYLATION (BY SIMILARITY).
SO	SEQUENCE	770 AA; 88039 MW; D74A0C76954754ED CRC64;

Query	March	Similarity	59.6%	Score	459	DB	1	length	770
Best	Local	Similarity	100.0%	Pred.	No.	0			
Matches	459	Conservative	0	Mismatches	0	Indels	0	Gaps	0

Qy	312	LFRNLKSAFVVEROPCPMPHDPRLVIRKTVQFTFKRLLYKPELNYOLKIKVICKD	371
Db	312	LFRNLKSAFVVEROPCPMPHDPRLVIRKTVQFTFKRLLYKPELNYOLKIKVICKD	371
Qy	372	SGDVAALGSKRFNLGLTNTRYKYNNEESNGLSASEFHLLTREORCGNGGRANDASLI	431
Db	372	SGDVAALGSKRFNLGLTNTRYKYNNEESNGLSASEFHLLTREORCGNGGRANDASLI	431
Qy	432	VTEELHLTFTEFEVYHOGKIDLETHSLPYYVYISNICOMPANMASILMYNMLTNPKVYN	491
Db	432	VTEELHLTFTEFEVYHOGKIDLETHSLPYYVYISNICOMPANMASILMYNMLTNPKVYN	491
Qy	492	FPTKPPICITWDOVAEVLISWOFSSSTTKRGLSIEQLTLTAEKLLGPGVNSGCOITWAKFC	551
Db	492	FPTKPPICITWDOVAEVLISWOFSSSTTKRGLSIEQLTLTAEKLLGPGVNSGCOITWAKFC	551
Qy	552	ENMACKGGSFVWVWLDNIIDLYKYYTLALMNGCYIMGFIISKERERAILSTRKPGFFLLRFS	611
Db	552	ENMACKGGSFVWVWLDNIIDLYKYYTLALMNGCYIMGFIISKERERAILSTRKPGFFLLRFS	611
Qy	612	ESSKSGGVTFWVWEKDIGSKTOIQSVEEYTRQOOLNNMSFAEIMGYKIMDATNIIIVSLV	671
Db	612	ESSKSGGVTFWVWEKDIGSKTOIQSVEEYTRQOOLNNMSFAEIMGYKIMDATNIIIVSLV	671
Qy	672	YLIPDIKPEEAFGKYCRPESOEHPADPGSAAPYLKTFICVTPPTCSNTIDLPMSPTLL	731
Db	672	YLIPDIKPEEAFGKYCRPESOEHPADPGSAAPYLKTFICVTPPTCSNTIDLPMSPTLL	731

RESULT	3		
STA3	HUMAN		
ID	STA3_HUMAN	STANDARD:	PRT: 770 AA.
AC	P40763.		
DT	01-FEB-1995 (Rel. 31, Created)		
DT	01-FEB-1995 (Rel. 31, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3 (ACUTE-PHASE RESPONSE FACTOR).		
DE	STAT3 OR APF.		
GN	STAT3		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Placenta;		
RX	MEDLINE=94208062; PubMed=7512451;		

RT pathway." www.ncbi.nlm.nih.gov/pubmed/7701321
 RL Cell 77:63-71(1994).
 RN (2)
 RP PHOSPHORYLATION ON SERINE.
 RX MEDLINE=95215843; PubMed=7701321;
 RA Zhang X., Fienis J., Li H.-C., Schindler C., Chen-Kiang S.;
 RT "Requirement for serine phosphorylation for formation of STAT-promoter
 RT complexes."
 RL Science 267:1990-1994(1995).
 CC -I- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6
 CC (IL-6)-RESPONSIVE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS
 CC ACUTE-PHASE PROTEIN GENES.
 CC -I- PATHWAY: INVOLVED IN THE G1/S-MEDIATED SIGNALING PATHWAY.
 CC -I- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
 CC MEMBER (AT LEAST STAT1).
 CC -I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION.
 CC -I- TISSUE SPECIFICITY: HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL
 CC MUSCLE, KIDNEY, AND PANCREAS.
 CC -I- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNF,
 CC LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION
 CC IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3
 CC HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY.
 CC -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -I- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L29277; AAA58374.1; -
 DR HSSP: P42224; 1BP5.
 DR TRANSFAC: T01493; -
 DR MIM: 102562; -
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR Transcription regulation; DNA-binding; Nuclear protein;
 KW Phosphorylation; SH2 domain.

```

FT DOMAIN 580 670 SH2.
FT MOD_RES 705 705 PHOSPHORYLATION (BY JAKS) (BY
FT MOD_RES 727 727 SIMILARITY).
FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
SO SEQUENCE 770 AA: 88052 MW: A3DCEBF15B3B5560 CRC64;

Query Match 28.7%; Score 221; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 5,1e-220;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 FPMELRQFLAPWIESQDMATYAAKSESHATLVFNHLGELIDQYTSRFLQESNVLQYHQLRR 85
DB 26 FPMELRQFLAPWIESQDMATYAAKSESHATLVFNHLGELIDQYTSRFLQESNVLQYHQLRR 85
OY 86 IKOFLQSRYLEKPMELIARIYARCLMEESRLQTAATAAGGGOANHPPTAAVTEKQOMLE 145
DB 86 IKOFLQSRYLEKPMELIARIYARCLMEESRLQTAATAAGGGOANHPPTAAVTEKQOMLE 145
OY 146 QHLDVKKRVQDLEQKKKKYVENLDQDFPNYKTKLSGDMQDLNGNOSTYRQKMOQLEQ 205
DB 146 QHLDVKKRVQDLEQKKKKYVENLDQDFPNYKTKLSGDMQDLNGNOSTYRQKMOQLEQ 205
OY 206 MTLALDQMRRSIVSELAGLSAMEYVQKTLTDEELADMKRR 246
DB 206 MTLALDQMRRSIVSELAGLSAMEYVQKTLTDEELADMKRR 246

RESULT 4
STAT_MOUSE STANDARD; PRT; 749 AA.
ID STAT_MOUSE
AC P42225;
DT 01-NOV-1995 (Rel. 32, Created)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.
GN STAT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=94255416; PubMed=7545930;
RA Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "Stat3 and Stat4: members of the family of signal transducers and
activators of transcription.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4806-4810(1994).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
ALPHA, STAT1-BETA, STAT2) OF ISGF3 BECOME PHOSPHORYLATED ON
TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT. IN
RESPONSE TO IFN GAMMA, STAT1 FORMS HOMODIMERS, THAT ALSO
TRANSLocate INTO THE NUCLEUS TO ACTIVATE IFN GAMMA-RESPONSIVE
GENES. INTERACTS WITH NMI (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLocATED INTO THE NUCLEUS IN
RESPONSE TO PHOSPHORYLATION.
CC -1- INDUCTION: BY IFN AND EGF.
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-GAMMA, IFN-ALPHA,
PDGF, AND EGF. SERINE PHOSPHORYLATION IS ALSO REQUIRED FOR MAXIMAL
TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U06924; AAA19454.1; -.
DR HSSP: P42224; 1BF5.
DR TRANSPAC: T01575; -.
DR MGD: MGI:103063; Stat1.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS20001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; SH2 domain;
KW Phosphorylation.
FT DOMAIN 573 670
FT MOD_RES 701 701 SH2.
FT MOD_RES 701 701 PHOSPHORYLATION (BY JAKS) (BY
FT MOD_RES 701 701 SIMILARITY).
SO SEQUENCE 749 AA: 87197 MW: 249D919952BE65F1 CRC64;

Query Match 1.6%; Score 12; DB 1; Length 749;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 585 IMGFISKEKERRA 596
DB 578 IMGFISKEKERRA 589

RESULT 5
STAT_HUMAN STANDARD; PRT; 750 AA.
ID STAT_HUMAN
AC P42224;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1-ALPHA/BETA
DE (TRANSCRIPTION FACTOR ISGF-3 COMPONENTS P91/P84).
GN STAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 514-524; 654-660 AND 667-672.
RX MEDLINE=92366557; PubMed=1502203;
RA Schindler C., Fu X.-Y., Imbrota T., Aebersold R., Darnell J.E. Jr.;
RT "Proteins of transcription factor ISGF-3: one gene encodes the 91-and
RT 84-kDa ISGF-3 proteins that are activated by interferon alpha.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7836-7839(1992).
RN [2]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RA Yan R., Qureshi S., Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in
RT coincident sites in Stat1 and Stat2.";
RL Nucleic Acids Res. 23:459-463(1995).
RN [3]
RP PHOSPHORYLATION OF TYR-701.
RX MEDLINE=95386533; PubMed=7657660;
RA Quella F.W., Thierfelder W., Wiltuhn B.A., Tang B., Cohen S.,
RA Igle J.N.;
RT "Phosphorylation and activation of the DNA binding activity of
RT purified Stat1 by the Janus protein-tyrosine kinases and the
RT epidermal growth factor receptor.";
RL J. Biol. Chem. 270:20775-20780(1995).
RN [4]
RP PHOSPHORYLATION OF SER-727, AND MUTAGENESIS.
RX MEDLINE=95354205; PubMed=7543024;
RA Wen Z., Zhong Z., Darnell J.E. Jr.;

```

RT "Maximal activation of transcription by Stat1 and Stat3 requires both
 RT tyrosine and serine phosphorylation."
 RL Cell 82:241-250(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 136-710.
 RX MEDLINE:98292180; PubMed:9630226;
 RA Chen X., Vinkemeier U., Zhao Y., Jeruzalmi D., Darnell J.E. Jr.,
 KA Kutlany J.,
 RT "Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to
 RT DNA."
 RL Cell 93:827-839(1998).
 CC -I- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
 CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
 CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
 CC -I- PATHWAY: IFN-SIGNALING PATHWAY.
 CC -I- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
 CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
 CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
 CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
 CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT. IN
 CC RESPONSE TO IFN GAMMA, STAT1 FORMS HOMODIMERS, THAT ALSO
 CC TRANSLOCATE INTO THE NUCLEUS TO ACTIVATE IFN GAMMA-RESPONSIVE
 CC GENES. INTERACTS WITH NFI.
 CC -I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION.
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS, ALPHA/P91 (SHOWN HERE) AND
 CC BETA/P84; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-GAMMA, IFN-ALPHA,
 CC PDGF, AND EGF. SERINE PHOSPHORYLATION IS ALSO REQUIRED FOR MAXIMAL
 CC TRANSCRIPTIONAL ACTIVITY (LACKING IN BETA FORM).
 CC -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -I- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M97935; AAB64012.1; -;
 DR EMBL: M97936; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18662; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18663; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18664; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18665; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18666; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18667; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18668; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18669; -; NOT_ANNOTATED_CDS.
 DR EMBL: U18670; -; NOT_ANNOTATED_CDS.
 DR PDB: 1BF5; 12-AUG-98.
 DR TRANSFAC: T01492; -;
 DR MIM: 600555; -;
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SMO0252; SH2; 1.
 DR PROSITE: PS0001; SH2; 1.
 DR Transcription regulation: DNA-binding; Nuclear protein;
 DR Phosphorylation; SH2 domain; Alternative splicing; 3D-structure.
 KM DOMAIN 573 670
 FT MOD_RES 701 701 PHOSPHORYLATION (BY JAKS).
 FT MOD_RES 727 727 PHOSPHORYLATION.
 FT VARSPIC 713 750 MISSING (IN ISOFORM BETA).
 FT MUTAGEN 727 727 S->A: DECREASED TRANSCRIPTIONAL
 FT ACTIVATION.
 FT CONFLICT 261 271 MISSING (IN M97936).
 FT SEQUENCE 750 AA; 87334 MW; 054A813522364BA6 CRC64;

Query Match 1.6%; Score 12; DB 1; Length 750;
 Best Local Similarity 100.0%; Pred. No. 0.00066;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 585 IMGFISKERERA 596
 DB 578 IMGFISKERERA 589
 |||||||
 RESULT 6
 STA4_HUMAN
 ID STA4_HUMAN STANDARD; PRI: 748 AA.
 AC 014765;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 4.
 GN STA4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xu X., Sun Y.L., Hoey T.;
 RT "The STAT amino-terminal domain mediates cooperative DNA binding
 RT and confers selective sequence recognition."
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
 CC ACTIVATION OF TRANSCRIPTION.
 CC -I- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
 CC MEMBER (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION.
 CC -I- PTM: TYROSINE PHOSPHORYLATED. SERINE PHOSPHORYLATION IS ALSO
 CC REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -I- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L78440; AAB05605.1; -;
 DR MIM: 600558; -;
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SMO0252; SH2; 1.
 DR PROSITE: PS0001; SH2; 1.
 DR Transcription regulation: DNA-binding; Nuclear protein;
 KM Phosphorylation; SH2 domain.
 KM DOMAIN 569 664
 FT MOD_RES 693 693 PHOSPHORYLATION (BY JAKS) (BY
 FT SIMILARITY).
 FT MOD_RES 721 721 PHOSPHORYLATION (BY SIMILARITY).
 FT SEQUENCE 748 AA; 85940 MW; 11E43803A9AFAFFA CRC64;
 QY 603 PGFILLRSES 613
 DB 592 PGFILLRSES 602
 Query Match 1.4%; Score 11; DB 1; Length 748;
 Best Local Similarity 100.0%; Pred. No. 0.0071;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 7
STA2_MOUSE STANDARD: PRT: 749 AA.
ID STA2_MOUSE
AC P42228:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 4.
GN STA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RA MEDLINE=94255416; PubMed=7545930;
RA Zhong Z., Wen Z., Darneil J.E. Jr.;
RA "Stat3 and Stat4: members of the family of signal transducers and
RT activators of transcription.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4806-4810(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Testis;
RX MEDLINE=94277038; PubMed=8007943;
RA Yamamoto K., Quelle F.W., Thierfelder W.E., Kreider B.L.,
RA Gilbert D.J., Jenkins N.A., Copeland N.G., Silvennoinen O.,
RA Thle J.N.;
RT "Stat4, a novel gamma interferon activation site-binding protein
RT expressed in early myeloid differentiation.";
RL Mol. Cell. Biol. 14:4342-4349(1994).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.45 ANGSTROMS) OF 1-123.
RX MEDLINE=98128033; PubMed=9461439;
RA Vinkemeier U., Moarefi I., Darneil J.E. Jr., Kuriyan J.;
RT "Structure of the amino-terminal protein interaction domain of
RT Stat-4.";
RL Science 279:1048-1052(1998).
CC -1- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- PTM: TYROSINE PHOSPHORYLATED. SERINE PHOSPHORYLATION IS ALSO
CC REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL: U06922; AAA19453.1; -
DR EMBL: U09351; AAA19692.1; -
DR PDB: 1BGF; 16-SEP-98.
DR TRANSFAC: T01576; -
DR MGD: MGI:103062; Stat4.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2_1.
DR Pfam: PF01017; STAT_1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
KM Transcription regulation: DNA-binding; Nuclear protein;
KM Phosphorylation: SH2 domain; 3D-structure.
FT MOD_RES 570 665 SH2
FT MOD_RES 694 694 PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).

```

```

FT MOD_RES 722 722 PHOSPHORYLATION (BY SIMILARITY).
FT CONFLICT 199 199 MISSING (IN REF. 2).
FT CONFLICT 638 638 A -> P (IN REF. 2).
SQ SEQUENCE 749 AA: 85940 MW: A88B837E49CFCEBC CRC64:

Query Match 1.4%; Score 11; DB 1; Length 749;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 603 PGFTLRFSES 613
Db 593 PGFTLRFSES 603

RESULT 8
STA2_HUMAN STANDARD: PRT: 851 AA.
ID STA2_HUMAN
AC P52630; Q16430; Q16431;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2 (P113).
GN STAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RA Yan R., Qureshi S., Zhong Z., Wen Z., Darneil J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in
RT coincident sites in Stat1 and Stat2.";
RL Nucleic Acids Res. 23:459-463(1995).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=9236558; PubMed=1502204;
RA Fu X.-Y., Schindler C., Impirota T., Aebersold R., Darneil J.E. Jr.;
RT "The proteins of ISGF-3, the interferon alpha-induced transcriptional
RT activator, define a gene family involved in signal transduction.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7840-7843(1992).
RN [3]
RP ALTERNATIVE SPLICING.
RX MEDLINE=96176320; PubMed=8601453;
RA Sugiyama T., Nishio Y., Kishimoto T., Akira S.;
RT "Identification of alternative splicing form of Stat2.";
RL FEBS Lett. 381:191-194(1996).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=97172544; PubMed=9020188;
RA Bluyssen H.A., Levy D.E.;
RT "Stat2 is a transcriptional activator that requires sequence-specific
RT contacts provided by stat1 and p48 for stable interaction with DNA.";
RL J. Biol. Chem. 272:4600-4605(1997).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
CC TOGETHER WITH ISGF3 GAMMA (P48). A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA.
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U18671: AAA98760.1: -
 DR EMBL: M97934: -; NOT_ANNOTATED_CDS.
 DR EMBL: S81491: AAB36226.1: -
 DR EMBL: S81491: AAB36227.1: ALT_SEQ.
 DR HSSP: P42224: 1BF5.
 DR MIM: 600556: -
 DR InterPro: IPR000980: SH2.
 DR InterPro: IPR001217: STAT.
 DR Pfam: PF00017: SH2; 1.
 DR Pfam: PF01017: STAT; 1.
 DR SMART: SM00252: SH2; 1.
 DR PROSITE: PS50001: SH2; 1.
 KM Transcription regulation; DNA-binding; Nuclear protein;
 KW Phosphorylation; SH2 domain; Alternative splicing.
 FT DOMAIN 572 667 SH2.
 FT MOD_RES 690 690 PHOSPHORYLATION (BY JAKS) (BY
 FT SIMILARITY).
 FT VARSPIC 621 652 DKVLIVSVPYREKVLQSLPLTEIRHYQLT -> GQLLC
 FT PAIPASPECLAPLPLCLPLASIALN (IN SHORT
 FT ISOFORM).
 FT VARSPIC 653 851 MISSING (IN SHORT ISOFORM).
 FT SEQUENCE 851 AA: 97916 MW: 9474674CB7A3215 CRC64;

Query Match 1.4%; Score 11; DB 1; Length 851;
 Best Local Similarity 100.0%; Pred. No. 0.0079;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 604 GTFILRFSESS 614
 DB 596 GTFILRFSESS 606
 |||||||

RESULT 9
 STA2_PIG STANDARD: PRT: 864 AA.
 AC 002799;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.
 GN STAT2
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle; N.A.
 RA Ito Y., Mikawa S., Kobayashi E., Wada Y., Minezawa M.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
 CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
 CC TRANSCRIPTION FACTOR IS TERMED ISGF3 (BY SIMILARITY).
 CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
 CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
 CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
 CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
 CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
 CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC This SWISS entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB004061: BAA20332.1: -
 DR InterPro: IPR000980: SH2.
 DR InterPro: IPR001217: STAT.
 DR Pfam: PF00017: SH2; 1.
 DR Pfam: PF01017: STAT; 1.
 DR SMART: SM00252: SH2; 1.
 DR PROSITE: PS50001: SH2; 1.
 KM Transcription regulation; DNA-binding; Nuclear protein;
 KW Phosphorylation; SH2 domain.
 FT DOMAIN 572 667 SH2.
 FT MOD_RES 690 690 PHOSPHORYLATION (BY JAKS) (BY
 FT SIMILARITY).
 FT SEQUENCE 864 AA: 99060 MW: 2f8CC95569FB5B20 CRC64;

Query Match 1.2%; Score 9; DB 1; Length 864;
 Best Local Similarity 100.0%; Pred. No. 0.94;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 604 GTFILRFSE 612
 DB 596 GTFILRFSE 604
 |||||||

RESULT 10
 STA2_MOUSE STANDARD: PRT: 923 AA.
 AC Q9WVZ2: Q64189; Q64250; Q64188;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.
 GN STAT2
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1;
 RA Paulson M.S., Mui A., Levy D.E.;
 RT "Molecular cloning and characterization of murine Stat2.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 595-658 FROM N.A. (ISOFORMS A AND B/C).
 RX MEDLINE=96176320; PubMed=8601453.
 RA Sugiyama T., Nishio Y., Kishimoto T., Akira S.;
 RL "Identification of alternative splicing form of Stat2.";
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
 CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
 CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
 CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
 CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
 CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
 CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
 CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B/C; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BRAIN, LUNG, HEART, SPLEEN,
 CC LIVER, KIDNEY, MUSCLE, AND THE TESTIS.
 CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA (BY

```

CC      1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC      -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF088862; AAD38329.1; -
DR      EMBL; S81342; AAB36228.2; -
DR      EMBL; S81342; AAB36231.1; -
DR      EMBL; S81342; AAB36230.1; ALT_SEQ.
DR      HSP; P42224; 1BF5.
DR      MCD; MGI:103039; Stat2.
DR      InterPro: IPR000980; SH2.
DR      InterPro: IPR001217; STAT.
DR      Pfam; PF00017; SH2; 1.
DR      Pfam; PF01017; STAT; 1.
DR      SMART; SM00252; SH2; 1.
DR      PROSITE; PS00001; SH2; 1.
KW      Transcription regulation; DNA-binding; Nuclear protein;
KW      phosphorylation; SH2 domain; Alternative splicing.
FT      DOMAIN 571 666
FT      MOD_RES 689 689
FT      PHOSPHORYLATION (BY JAKS) (BY
FT      SIMILARITY).
FT      VARSPIC 620 643
FT      HKEIYSQPYTKVELSLPLEI -> GQHPPYHSCSL
FT      SARHPITRLP (IN SHORT ISOFORM).
FT      VARSPIC 644 923
FT      MISSING (IN SHORT ISOFORM).
FT      CONFLICT 596 596
FT      T -> A (IN REF. 2).
FT      CONFLICT 620 620
FT      H -> D (IN REF. 2).
SQ      SEQUENCE 923 AA: 105416 MW: D50BB54C35B0774 CRC64;

Query Match 1.2%; Score 9; DB 1; Length 923;
Best Local Similarity 100.0%; Pred. No. 0.99;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 604 GTFLLRFSE 612
DB 595 GTFLLRFSE 603

RESULT 11
HDNO_ARTOX STANDARD: PRT: 458 AA.
AC P08159;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 6-HYDROXY-D-NICOTINE OXIDASE (EC 1.5.3.6) (6-HDNO).
OS Arthrobacter oxidans.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococcales; Micrococcaceae; Arthrobacter.
OC NCBI_TaxID=1671;
RN RP SEQUENCE FROM N.A.
RX MEDLINE=87304263; PubMed=3622516;
RA Brandesch R., Hinkkanen A.E., Mauch L., Nagursky H., Decker K.;
RT "6-Hydroxy-D-nicotine oxidase of Arthrobacter oxidans. Gene structure
RT of the flavoenzyme and its relationship to 6-hydroxy-L-nicotine
RT oxidase."
RL Eur. J. Biochem. 167:315-320(1987).
RN [2]
RP REVISIONS.
RA Brandesch R.;
RX Submitted (SEP-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP MUTAGENESIS OF HIS-71.
RX MEDLINE=90033359; PubMed=2680607;

```

```

RA Mauch L., Biehler V., Brandesch R.;
RT "Site-directed mutagenesis of the FAD-binding histidine of
RT 6-hydroxy-D-nicotine oxidase. Consequences on flavinylation and
RT enzyme activity."
RL FEBS Lett. 257:86-88(1989).
RN [4]
RP MUTAGENESIS OF ARG-67 AND SER-68.
RX MEDLINE=90330600; PubMed=2115879;
RA Mauch L., Biehler V., Brandesch R.;
RT "Lysine can replace arginine 67 in the mediation of covalent
RT attachment of FAD to histidine 71 of 6-hydroxy-D-nicotine oxidase."
RL J. Biol. Chem. 265:12761-12762(1990).
CC -1- CATALYTIC ACTIVITY: (D)-6-HYDROXYNICOTINE + H(2)O + O(2) -
CC 1-(6-HYDROXY)PYRID-3-YL)-4-(METHYLAMINO)BUTAN-1-ONE + H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: DEGRADATION OF NICOTINE (INDUCIBLE).
CC -1- SIMILARITY: TO OTHER OXYGEN OXIDOREDUCTASES THAT COVALENTLY BIND
CC FAD.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR      EMBL; X05999; CAA29416.1; -
DR      PIR; S00087; DEIOHN.
DR      InterPro: IPR001575; Oxid_FAD_bind.
DR      Pfam; PF01565; FAD_binding_4; 1.
DR      PROSITE; PS00862; OX2_COVAL_FAD; 1.
KW      Oxidoreductase; Flavoprotein; FAD.
FT      BINDING 71 71
FT      MUTAGEN 67 67
FT      MUTAGEN 67 67
FT      R -> A: NO FAD INCORPORATION.
FT      R -> K: NO EFFECT ON FAD INCORPORATION, BUT
FT      REDUCED ACTIVITY.
FT      S -> A: NO EFFECT ON FAD INCORPORATION OR
FT      ON ACTIVITY.
FT      H -> C,Y,S: NO FAD INCORPORATION, ABOLISH
FT      OR DIMINISH SIGNIFICANTLY THE ACTIVITY.
SQ      SEQUENCE 458 AA: 48786 MW: 6783E44D66DC841 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 458;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 VESLAGLL 225
DB 219 VESLAGLL 226

RESULT 12
MNEI_YEAST STANDARD: PRT: 663 AA.
AC P24720;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MNEI PROTEIN.
DE MNEI OR MNE OR YOR350C OR O6353.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN RP SEQUENCE FROM N.A.
RX STRAIN=S288C / FY73;
RX MEDLINE=97103776; PubMed=8948102;
RA Purnelle B., Goffeau A.;
RT "Nucleotide sequence analysis of a 40 kb segment on the right arm of
RT yeast chromsome XV reveals 18 open reading frames including a new

```

RT pyruvate kinase and three homologues to chromosome I genes.*;
RL Yeast 12:1475-1481(1996).
[2]
RP SEQUENCE OF 1-219 FROM N.A.
RX MEDLINE-92158649: PubMed-1741279;
RA Ieem S.H., Ogawa H.;
RT "The MRE4 gene encodes a novel protein kinase homologue required for
RT meiotic recombination in Saccharomyces cerevisiae.";
RL Nucleic Acids Res. 20:449-457(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X95720; CAA65037.1; -
DR EMBL: X75258; CAA99678.1; -
DR EMBL: X63112; CAA44826.1; -
DR PIR: S19073; S19073.
DR PIR: S20175; S20175.
DR SGD: S0005877; MNE1.
SQ SEQUENCE 663 AA; 77229 MW; 208DD629E7034858 CRC64;

OY 590 SKERERAI 597
Db 249 SKERERAI 256

RESULT 13
STAT_DROME STANDARD; PRT; 761 AA.
AC 024151: 024181: 09VDL8:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND TRANSCRIPTION ACTIVATOR (MARELLE PROTEIN)
DE (D-STAT).
GN STAT92E OR MARE OR MRL OR STAT OR CG4257.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RC TISSUE=Embryo;
RX MEDLINE-96190808: PubMed-8608595;
RA Hou X.S., Meinick M.B., Perrimon N.;
RT "Marelle acts downstream of the Drosophila HOP/JAK kinase and encodes
RT a protein similar to the mammalian STATs.";
RL Cell 84:411-419(1996).
RN [2]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Embryo;
RX MEDLINE-96190809: PubMed-8608596;
RA Yan R., Small S., Desplan C., Dearolf C.R., Darnell J.E. Jr.;
RT "Identification of a Stat gene that functions in Drosophila
RT development.";
RL Cell 84:421-430(1996).
RN [3]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RC STRAIN=BERKELEY;
RX MEDLINE-20196006: PubMed-10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Abmayri A., An H.-J., Andrews-Pfennoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benus P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Daynoff L.B., Davies P.,
RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertler S., Fletcher W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hardin N.L., Harvey D., Heiman T.J., Hernandez J.R., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhen M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong M., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: SIGNAL TRANSDUCTION AND ACTIVATION OF TRANSCRIPTION.
CC EARLY EMBRYO BY ACTIVATING SPECIFIC STRIPS OF PAIR RULE GENE
CC EXPRESSION. THE JAKUS KINASE-STAT PATHWAY IS CONNECTED TO
CC DROSOPHILA EARLY DEVELOPMENT. MUTANTS EXHIBIT ABERRANT EXPRESSION
CC OF THE PAIR RULE GENE EVEN-SKIPPED AT THE CELLULAR BLASTODERM
CC STAGE, LEADING TO LARVAL SEGMENTATION DEFECTS.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND
CC A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT EMBRYONIC, LARVAL,
CC PUPAL, AND ADULT STAGES, WITH SOME DECREASE IN THE LATE EMBRYONIC
CC STAGES. THE EXPRESSION IS UNIFORM IN UNFERTILIZED OR JUST
CC FERTILIZED EGGS, SUGGESTING MATERNALLY DEPOSITED mRNA. AT
CC BLASTODERM STAGE, EXPRESSION PATTERN SHOWS STRIPES, THAT ARE
CC REMINISCENT OF MANY PAIR RULE GENES PATTERN.
CC -1- PTM: TYROSINE PHOSPHORYLATED BY HOSPOCCH. PHOSPHORYLATION IS
CC REQUIRED FOR DNA-BINDING ACTIVITY AND DIMERIZATION.
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC IN PARTICULAR TO MAMMALIAN STATs AND STATc.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U40070; AAC46984.1; -
DR EMBL: U46688; AAB02195.1; -
DR EMBL: AE003731; AAF55773.1; -
DR HSSP: P42224; 1BF5.

```

DR FlyBase; FBgn0016917; Stat92E.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF01017; STAT; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; SH2 domain;
KW Phosphorylation; Alternative splicing; Developmental protein.
FT DOMAIN 594 658 SH2.
FT MOD_RES 711 711 PHOSPHORYLATION (BY JAKS).
FT VARSPPLIC 699 705 MISSING (IN SHORT ISOFORM).
FT CONFLICT 105 105 S -> T (IN REF. 2).
FT CONFLICT 648 648 L -> H (IN REF. 2).
SQ SEQUENCE 761 AA; 86414 MW; BF3A622A29899161 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 761;
Best Local Similarity 100.0%; Pred. No. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
DB 620 GTFLLRFS 627

RESULT 14
ST5B_HUMAN STANDARD; PRT; 786 AA.
ID ST5B_HUMAN
AC P51692.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5B.
GN STAT5B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96311205; PubMed=87332682;
RA Shiva C.M., Lu H., Day R.N.;
RT "Characterization and cloning of STAT5 from IM-9 cells and its
RL Mol. Endocrinol. 10:508-518(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96210005; PubMed=8631883;
RA Lin J.-X., Metz J., Modi W.S., John S., Leonard W.J.;
RT "Cloning of human Stat5B. Reconstitution of interleukin-2-induced
RT Stat5A and Stat5B DNA binding activity in COS-7 cells.";
RL J. Biol. Chem. 271:10738-10744(1996).
RN [3]
RP INTERACTION WITH NMI.
RX PubMed=9989503;
RA Zhu M.-H., John S., Berg M., Leonard W.J.;
RT "Functional association of Nmi with Stat5 and Stat1 in IL-2- and
RT IFNgamma-mediated signaling.";
RL Cell 96:121-130(1999).
CC -1- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION. BINDS TO THE GAS ELEMENT AND
CC ACTIVATES PRL-INDUCED TRANSCRIPTION.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY). INTERACTS WITH NMI.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -1- PTM: TYROSINE PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U48730; AAC50485.1; -.
DR EMBL; U47686; AAC50491.1; -.
DR HSSP; P42224; 1BF5.
DR MIM; 604260; -.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; STAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT; 1.
DR SMART; SMC_52; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 589 686 SH2.
FT MOD_RES 699 699 PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).
FT CONFLICT 230 231 A -> P (IN REF. 2).
FT CONFLICT 628 628 S -> T (IN REF. 2).
FT CONFLICT 717 717 D -> DA (IN REF. 2).
FT CONFLICT 720 720 R -> G (IN REF. 2).
SQ SEQUENCE 786 AA; 89880 MW; E4265C4C8EB824B6 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 786;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
DB 613 GTFLLRFS 620

RESULT 15
ST5B_MOUSE STANDARD; PRT; 786 AA.
ID ST5B_MOUSE
AC P42232; O60804;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5B.
GN STAT5B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6 X A/J; TISSUE=liver;
RX MEDLINE=95237198; PubMed=7720707;
RA Mui A.L.-F., Wakao H., O'Farrell A.-M., Harada N., Miyajima A.;
RT "Interleukin-3, granulocyte-macrophage colony stimulating factor and
RT interleukin-5 transduce signals through two STAT5 homologs.";
RL EMBO J. 14:1166-1175(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=96004632; PubMed=7568026;
RA Liu X., Robinson G.W., Gouilleux F., Groner B., Hennighausen L.;
RT "Cloning and expression of Stat5 and an additional homologue (Stat5b)
RT involved in prolactin signal transduction in mouse mammary tissue.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:8831-8835(1995).
CC -1- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION. BINDS TO THE GAS ELEMENT AND
CC ACTIVATES PRL-INDUCED TRANSCRIPTION.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -----

```

```
CC -1- TISSUE SPECIFICITY: IN THE VIRGIN, FOUND IN MOST TISSUES.
CC PARTICULARLY ABUNDANT IN MUSCLE TISSUE OF VIRGIN AND LACTATING
CC FEMALES, AND OF MALES.
CC -1- DEVELOPMENTAL STAGE: DETECTED BOTH IN VIRGIN MOUSE AND AFTER
CC MAMMARY GLAND INVOLUTION. THE LEVEL OF STAT5A INCREASES CONSTANTLY
CC DURING PREGNANCY, BUT DECREASES DURING LACTATION.
CC -1- PTM: TYROSINE PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z48539; CAA88420.1; -.
DR EMBL: U21110; AAC52282.1; -.
DR HSSP: P42224; 1BF5.
DR TRANSFAC: T00944; -.
DR MGD: MGI:103035; Stat5b.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
DR Transcription regulation: DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 589 686
FT MOD_RES 699 699
FT PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).
FT CONFLICT 433 433 E->G (IN REF. 2).
FT SEQUENCE 786 AA: 90002 MW: A8FE76405E41B2EF CRC64;
```

Query Match 1.0%; Score 8; DB 1; Length 786;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 604 GTFLLRFS 611
|||||
Db 613 GTFLLRFS 620

Search completed: March 19, 2002, 15:27:00
Job time: 252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2002, 14:00:57 ; Search time 2747.57 Seconds

(without alignments)
17226.279 Million cell updates/sec

Title: US-08-212-185-11

Perfect score: 2869

Sequence: 1 GCCGCGACCGACGCGCGG.....AATTAAAAAAAAAAAAA 2869

Scoring table: IDENTITY_NUC

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenEmbl:*

1: gb_da:*

2: gb_hlg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pi:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_da:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_om:*

20: em_or:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_sy:*

28: em_un:*

29: em_vl:*

30: em_htgo_hum:*

31: em_htgo_inv:*

32: em_htgo_rod:*

33: em_htg_hum:*

34: em_htg_inv:*

35: em_htg_rod:*

36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2869	100.0	2869	6	AR083251	AR083251 Sequence
2	2869	100.0	2869	6	187790	187790 Sequence 11
3	2864.2	99.8	2869	6	AR121061	AR121061 Sequence
4	2864.2	99.8	2869	10	MMU06922	MMU06922 Mus musculus
5	2861	99.7	2869	6	ARI53599	ARI53599 Sequence
6	2767.8	96.5	2964	10	BC003806	BC003806 Mus muscu
7	2711.2	94.5	2924	10	MMU03709	MMU03709 Mus muscu
8	2705.2	94.3	2924	10	MMU08378	MMU08378 Mus muscu
9	2415.6	84.2	2652	6	188783	188783 Sequence 7
10	2415.6	84.2	2652	6	188784	188784 Sequence 8
11	2415.6	84.2	2652	10	MUSABRF	L29278 Mus musculu
12	2415.6	84.2	2652	22	E10408	E10408 CDNA encodi
13	2296	80.0	2924	10	RNSTRT3PT	X91810 R.norvegicu
14	2294	80.0	2310	6	188782	188782 Sequence 6
15	2144.2	74.7	2787	6	AR120980	AR120980 Sequence
16	2144.2	74.7	2787	6	AR121650	AR121650 Sequence
17	2144.2	74.7	2787	6	188780	188780 Sequence 3
18	2144.2	74.7	2787	6	188781	188781 Sequence 4
19	2144.2	74.7	2787	9	HUMAPRF	L29277 Homo sapien
20	2144.2	74.7	2787	22	E10407	E10407 CDNA encodi
21	2071.8	72.2	2663	9	BC000627	BC000627 Homo sapi
22	2026.8	70.6	2310	6	188779	188779 Sequence 2
23	2012.6	70.1	2344	6	A98087	A98087 Sequence 1
24	2012.6	70.1	2344	6	A98229	A98229 Sequence 1
25	1993	69.5	2313	5	HSN012463	A1012463 Homo sapi
26	1528.8	53.3	2653	5	AB017701	AB017701 Xenopus 1
27	1389.8	48.4	3071	5	DRE5693	AJ005693 Dario rer
28	1383	48.2	2304	5	OMU60333	BC004808 Mus muscu
29	698.6	24.3	2643	10	BC004808	AR083249 Sequence
30	682	23.8	2277	6	187788	187788 Sequence 7
31	682	23.8	2277	10	MMU06924	U06924 Mus musculu
32	682	23.8	2277	10	AF205604	AF205604 Rattus no
33	676	23.6	2260	10	AR083248	AR083248 Sequence
34	672.6	23.4	2607	6	187787	187787 Sequence 5
35	672.6	23.4	2607	6	HUMISGF3B	M97936 Human trans
36	672.6	23.4	2607	6	BC002704	BC002704 Homo sapi
37	672.6	23.4	2638	6	AR083247	AR083247 Sequence
38	672.6	23.4	3943	6	187786	187786 Sequence 3
39	672.6	23.4	3943	6	AR121648	AR121648 Sequence
40	672.6	23.4	4003	6	HUMISGF3A	M97935 Homo sapien
41	672.6	23.4	3943	6	ARI53600	ARI53600 Sequence
42	666.2	23.2	2265	5	OMU60331	U60332 Oncorhynch
43	661.4	23.1	2265	5	OMU60332	U60332 Oncorhynch
44	658.8	23.0	2265	5	MMU06923	U06923 Mus musculu
45	580.4	20.2	2533	10	MMU06923	U06923 Mus musculu

ALIGNMENTS

RESULT 1

LOCUS AR083251 2869 bp DNA

DEFINITION Sequence 11 from patent US 5976835.

ACCESSION AR083251

VERSION AR083251.1 GI:10010041

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2869)

AUTHORS Darnell,J.E., Jr., Schindler,C.W., Fu,X., Wen,Z., and Zhong,Z.

TITLE Nucleic acids encoding receptor recognition factor Stat1.alpha. and Stat1.beta. and methods of use thereof

JOURNAL Patent: US 5976835-A 11 02-NOV-1999;

FEATURES

source Location/Qualifiers

1..2869

/organism="unknown"

BASE COUNT 755 a 743 c 772 g 599 t

ORIGIN

OY 2101 CCAAGAGGAGGACATTGGAAGTACTGAGCCCGAGAGCCAGAGACCCGGAAGCCG 2160
|||||
Db 2101 CCAAGAGGAGGACATTGGAAGTACTGAGCCCGAGAGCCAGAGACCCGGAAGCCG 2160
OY 2161 ACCCAGGTGTCGTGCGCCGTACCTGGAAGCAAGTTCATCTGTGACACCAACCACT 2220
|||||
Db 2161 ACCCAGGTGTCGTGCGCCGTACCTGGAAGCAAGTTCATCTGTGACACCAACCACT 2220
OY 2221 GCAGCAATACCATTTGACCTGCGCATGTCCCGCCGACCTTTAGATTCATTGATGATG 2280
|||||
Db 2221 GCAGCAATACCATTTGACCTGCGCATGTCCCGCCGACCTTTAGATTCATTGATGATG 2280
OY 2281 GAAATTAACGGTGAAGTGTCTGAGCCCTGACGAGAGGAGCAATTGTGATCGCTACGTTG 2340
|||||
Db 2281 GAAATTAACGGTGAAGTGTCTGAGCCCTGACGAGAGGAGCAATTGTGATCGCTACGTTG 2340
OY 2341 ACATGATGTCGACCTGCGAGTGTGCTACCTCCCGCATGTGAGAGGTCGAAACAGAGCT 2400
|||||
Db 2341 ACATGATGTCGACCTGCGAGTGTGCTACCTCCCGCATGTGAGAGGTCGAAACAGAGCT 2400
OY 2401 GCAGAGACGTCGACCTGAGACACCTGCGCGTGCCTCCACCCCTAAGCAGCCGAAACCCATA 2460
|||||
Db 2401 GCAGAGACGTCGACCTGAGACACCTGCGCGTGCCTCCACCCCTAAGCAGCCGAAACCCATA 2460
OY 2461 TCGTCTGAACCTCTTACCTTTGTGCTTCCAGATTTTTTTTTTAATTTCTTACTTCTGCT 2520
|||||
Db 2461 TCGTCTGAACCTCTTACCTTTGTGCTTCCAGATTTTTTTTTTAATTTCTTACTTCTGCT 2520
OY 2521 ATCTTTGGCAATCTGGGCACTTTTAAAAAGAGAAATGAGTGTGAGTGTGATTAAC 2580
|||||
Db 2521 ATCTTTGGCAATCTGGGCACTTTTAAAAAGAGAAATGAGTGTGAGTGTGATTAAC 2580
OY 2581 TGTATGTAAAGAGAGACCTCTGAGTCTGGGATGGGGCTGAGAGCAGAAAGAGGC 2640
|||||
Db 2581 TGTATGTAAAGAGAGACCTCTGAGTCTGGGATGGGGCTGAGAGCAGAAAGAGGC 2640
OY 2641 AAGGGGAAACACCTCTGCTGCTGCGCGCTGCGCTCTTTTTCAGAGAGCTGGGGGTTGG 2700
|||||
Db 2641 AAGGGGAAACACCTCTGCTGCTGCGCGCTGCGCTCTTTTTCAGAGAGCTGGGGGTTGG 2700
OY 2701 TTGTTAGACAACTGCTCTGCTGCTGCGCATGCTGCTGTTGCCCACTCTGTGAGTGAT 2760
|||||
Db 2701 TTGTTAGACAACTGCTCTGCTGCTGCGCATGCTGCTGTTGCCCACTCTGTGAGTGAT 2760
OY 2761 ACCCATTCCTGGGAACTCCTGCTGCTGCGACTTTCACACTTGTCAATATCCACATAGAAC 2820
|||||
Db 2761 ACCCATTCCTGGGAACTCCTGCTGCTGCGACTTTCACACTTGTCAATATCCACATAGAAC 2820
OY 2821 TAGAGCTAAGCCAGAGAGGTTCTTTAAATTAATTAATTAATTAATTAATTAATTAAT 2869
|||||
Db 2821 TAGAGCTAAGCCAGAGAGGTTCTTTAAATTAATTAATTAATTAATTAATTAATTAAT 2869

RESULT 2
LOCUS 187790 2869 bp DNA PAT 10-AUG-1998
DEFINITION Sequence 11 from patent US 5716622.
ACCESSION 187790
VERSION 187790.1 GI:3407730
KEYWORDS
ORGANISM Unknown.
SOURCE Unknown.
REFERENCE 1 (bases 1 to 2869)
AUTHORS Darnell,J.E. Jr., Wen,Z., Horvath,C.M. and Zhong,Z.
TITLE Functionally active regions of signal transducer and activators of
transcription
Patent: US 5716622-A 11 10-FEB-1998;
JOURNAL Location/Qualifiers
FEATURES
1. 2869
BASE COUNT 755 a 743 c 772 g 599 t
ORIGIN

Query Match 100.0%: Score 2869: DB 6: Length 2869;
Best Local Similarity 100.0%: Pred No. 0;
Matches 2869: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCCGCACAGCCAGCCGCGCTCAGCTGCGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT 60
|||||
Db 1 GCCCGCACAGCCAGCCGCGGCTCAGCTGCGGCTCAGCCCGGAGAGTGCAGACCCCTGACT 60
OY 61 GCAGCAGATGAGCTCATGATGGAACCACTGTCAGCAGCTGTCAGACACCTACTTCAAGCAGC 120
|||||
Db 61 GCAGCAGATGAGCTCATGATGGAACCACTGTCAGCAGCTGTCAGACACCTACTTCAAGCAGC 120
OY 121 TGACACGCTGTACAGAGACAGCTTCCCATGAGCTGCGGCAATTTCTTGACACTTTGGA 180
|||||
Db 121 TGACACGCTGTACAGAGACAGCTTCCCATGAGCTGCGGCAATTTCTTGACACTTTGGA 180
OY 181 TTGAGAGTCAAGACTGGGCAATATGACAGCCAGCAAGAGTCAATGCCAGCTTGGTTTC 240
|||||
Db 181 TTGAGAGTCAAGACTGGGCAATATGACAGCCAGCAAGAGTCAATGCCAGCTTGGTTTC 240
OY 241 ATTAATCTTTGGGTGAATGATGACCAATATATAGCCGATTTCTGCAAGAGATCCAAATGTC 300
|||||
Db 241 ATTAATCTTTGGGTGAATGATGACCAATATATAGCCGATTTCTGCAAGAGATCCAAATGTC 300
OY 301 TCTATCGACCAACCTTGAAGAAATCAAGCAGTTTCTGACAGCAGGATCTTGAGAAAGC 360
|||||
Db 301 TCTATCGACCAACCTTGAAGAAATCAAGCAGTTTCTGACAGCAGGATCTTGAGAAAGC 360
OY 361 CAATGCAATTTGCCCGGATCTGCGCCGATGCTGTGGGAAAGTCTGCGCTCTCCAGA 420
|||||
Db 361 CAATGCAATTTGCCCGGATCTGCGCCGATGCTGTGGGAAAGTCTGCGCTCTCCAGA 420
OY 421 CGGAGCGACAGGAGCCAGCCAGAGGGGGCCAGGCCAACCCCAACCGCCGCTGAGTGA 480
|||||
Db 421 CGGAGCGACAGGAGCCAGCCAGAGGGGGCCAGGCCAACCCCAACCGCCGCTGAGTGA 480
OY 481 CAGAGAGCAGCAGATGTTGGAGCAGCATCTTGAGATGTCGCGAAACGAGTGCAGAGATC 540
|||||
Db 481 CAGAGAGCAGCAGATGTTGGAGCAGCATCTTGAGATGTCGCGAAACGAGTGCAGAGATC 540
OY 541 TGAACGAAATGAAGCTGTGGAGAACCTCCAGAGCACTTTGATTTAACTACAAAA 600
|||||
Db 541 TGAACGAAATGAAGCTGTGGAGAACCTCCAGAGCACTTTGATTTAACTACAAAA 600
OY 601 CCCTCAAGAGCCAGAGACATGACAGATCTGAATGGAACCAACCACTGTGACCCAGAC 660
|||||
Db 601 CCCTCAAGAGCCAGAGACATGACAGATCTGAATGGAACCAACCACTGTGACCCAGAC 660
OY 661 AGAAGATGACAGCAGCTGGAACAGATCTCACAGCCCTGACCAAGATGCGGAGAGCATTTG 720
|||||
Db 661 AGAAGATGACAGCAGCTGGAACAGATCTCACAGCCCTGACCAAGATGCGGAGAGCATTTG 720
OY 721 TGAGTAGCTGGGGGGGCTTTGTCAAGCAATGAGTACGTGCAGAAAGACACTACTGATG 780
|||||
Db 721 TGAGTAGCTGGGGGGGCTTTGTCAAGCAATGAGTACGTGCAGAAAGACACTACTGATG 780
OY 781 AAGAGCTGGCTGACTGGAAGAGCGGCCAGAGATGCGTGCATGCGAGGGCCCTCCCAACA 840
|||||
Db 781 AAGAGCTGGCTGACTGGAAGAGCGGCCAGAGATGCGTGCATGCGAGGGCCCTCCCAACA 840
OY 841 TCTGCTTGACCGCTGTGAAAACTGATTAATTCATTAGCAGATCTCAACTTCAGACCC 900
|||||
Db 841 TCTGCTTGACCGCTGTGAAAACTGATTAATTCATTAGCAGATCTCAACTTCAGACCC 900
OY 901 GCCAACAAATTAAGAACTGAGAGAGCTTCCAGAGAAATGTCTCAAGAGGCGACCCCTA 960
|||||
Db 901 GCCAACAAATTAAGAACTGAGAGAGCTTCCAGAGAAATGTCTCAAGAGGCGACCCCTA 960
OY 961 TCGTGCAGCAGCGCCCATGCTGTGAGAGAGAGATCTGAGAGCTGTTCAAGAACTTAATGA 1020
|||||
Db 961 TCGTGCAGCAGCGCCCATGCTGTGAGAGAGAGATCTGAGAGCTGTTCAAGAACTTAATGA 1020

QY	1021	AGA	GTCCTTCGTGGTGGAGGGCAGCCCTGCATG	CCCATGCA	CCCGAGCGCCCTTAG	1080		
Db	1021	AGAG	GGCCCTTCGTGGTGGAGGGCAGCCCTGCATG	CCCATGCA	CCCGAGCGCCCTTAG	1080		
QY	1081	TCAT	CATCAACACTGGGTGCAGTTTACCGCAAGAA	ATCGTGGTCAAA	TTTCTCTAGT	1140		
Db	1081	TCAT	CATCAACACTGGGTGCAGTTTACCGCAAGAA	ATCGTGGTCAAA	TTTCTCTAGT	1140		
QY	1141	TGA	ATTATCAGCTTTAAATTTAAAGTGTGAT	TAAGAC	TCGTGGGATGTTGCTGCC	1200		
Db	1141	TGA	ATTATCAGCTTTAAATTTAAAGTGTGAT	TAAGAC	TCGTGGGATGTTGCTGCC	1200		
QY	1201	TCAG	GGGTCTCGGAATTTAACATCTGGGCGCAAC	ACAAAGTGA	MAATGMACTGGAG	1260		
Db	1201	TCAG	GGGTCTCGGAATTTAACATCTGGGCGCAAC	ACAAAGTGA	MAATGMACTGGAG	1260		
QY	1261	AGT	CTAACACGCGAGCCCTGTCTGCAGAGTTC	CAAGCACTGAC	CCCTTAAGGAGCAGAA	1320		
Db	1261	AGT	CTAACACGCGAGCCCTGTCTGCAGAGTTC	CAAGCACTGAC	CCCTTAAGGAGCAGAA	1320		
QY	1321	GTG	GAATGGAGGCCGTGCCAATTTGTATCCTCT	TGATCGT	GACTGAGAGCTGCAC	1380		
Db	1321	GTG	GAATGGAGGCCGTGCCAATTTGTATCCTCT	TGATCGT	GACTGAGAGCTGCAC	1380		
QY	1381	TGAT	ACCTTGAACACTGAGAGTGTATACCAAG	CGCTCAAGATTGAC	CTAGACGACCCACT	1440		
Db	1381	TGAT	ACCTTGAACACTGAGAGTGTATACCAAG	CGCTCAAGATTGAC	CTAGACGACCCACT	1440		
QY	1441	CC	TTCGCAAGTTGTGTATCTCCACATCTGT	GCAGATGCC	AAATGCTTGGGCATCA	1500		
Db	1441	CC	TTCGCAAGTTGTGTATCTCCACATCTGT	GCAGATGCC	AAATGCTTGGGCATCA	1500		
QY	1501	TGT	GTATTAACATCTGATCCAAATTAACCCC	AAGACGTGA	ACTTTCAC	TAAAGCGGCAA	1560	
Db	1501	TGT	GTATTAACATCTGATCCAAATTAACCCC	AAGACGTGA	ACTTTCAC	TAAAGCGGCAA	1560	
QY	1561	TTG	GAACCTGGGACCAAGTGGCGGAGTGCT	CGCTGCGG	AGTTCTCTGCACACCA	CAAGC	1620	
Db	1561	TTG	GAACCTGGGACCAAGTGGCGGAGTGCT	CGCTGCGG	AGTTCTCTGCACACCA	CAAGC	1620	
QY	1621	GAG	GGCTGAGCATGAGACACTGACAACG	CTGGCTGAGA	AGCTCCTAG	GGCCTGGTGA	1680	
Db	1621	GAG	GGCTGAGCATGAGACACTGACAACG	CTGGCTGAGA	AGCTCCTAG	GGCCTGGTGA	1680	
QY	1681	ACT	ACTCAGGGTGTGCATACATGAGCTTAA	ATTGCA	AAAGAAAAA	ATGCTGGCAAG	1740	
Db	1681	ACT	ACTCAGGGTGTGCATACATGAGCTTAA	ATTGCA	AAAGAAAAA	ATGCTGGCAAG	1740	
QY	1741	GCT	TCTCTCTTCTGGGTCTGGCTAGACAA	TATCATG	CACTTGTG	AAAAAGTATATCTTGG	1800	
Db	1741	GCT	TCTCTCTCTTCTGGGTCTGGCTAGACAA	TATCATG	CACTTGTG	AAAAAGTATATCTTGG	1800	
QY	1801	CC	TTTGAATGAAGGCTATCATGATGGGTT	CTATCAG	CAAGAGCGGGAGCGGGCAT	CTC	1860	
Db	1801	CC	TTTGAATGAAGGCTATCATGATGGGTT	CTATCAG	CAAGAGCGGGAGCGGGCAT	CTC	1860	
QY	1861	TAA	CACAAAGCCCGCGGCACCTTCTCTAC	TGCGGTT	CAGAGACGAA	GAAGAG	1920	
Db	1861	TAA	CACAAAGCCCGCGGCACCTTCTCTAC	TGCGGTT	CAGAGACGAA	GAAGAG	1920	
QY	1921	GGG	CATCTTACCTGGGTGGTGAAGACAT	CGTGGCAAG	CCAGATCC	AGTCTGAG	1980	
Db	1921	GGG	CATCTTACCTGGGTGGTGAAGACAT	CGTGGCAAG	CCAGATCC	AGTCTGAG	1980	
QY	1981	AGC	CAACCAAGCAGACGCTGTAACAGAT	GTCA	TTTGTCTG	AAATCATCATG	GGCTATA	2040
Db	1981	AGC	CAACCAAGCAGACGCTGTAACAGAT	GTCA	TTTGTCTG	AAATCATCATG	GGCTATA	2040
QY	2041	AG	ATCATGATGGAGACCAATCTCTGTCT	CACTTGT	CTACTCT	TAACCCG	CATTC	2100
Db	2041	AG	ATCATGATGGAGACCAATCTCTGTCT	CACTTGT	CTACTCT	TAACCCG	CATTC	2100

QY	2101	CCAAGGAGGAGCGATTGGAAACTACTGTAAGCCGAGAGCCAGGACACCCCGAAAGCG	2160
Db	2101	CCAAAGGAGGAGCGATTGGAAAGTACTGTAGGCCGAGAGGCCAGGACACCCCGAAAGCG	2160
QY	2161	ACCCAGGTAAGTGTGCCCCGTACCTGAAGACCAAGTTCACTGTGTGACACCAACGACCT	2220
Db	2161	ACCCAGGTAAGTGTGCCCCGTACCTGAAGACCAAGTTCACTGTGTGACACCAACGACCT	2220
QY	2221	GCAGCAATACCATTTGACCTCCGATGTCCCCCGGACCTTTACATTCATTGATGACAGTTG	2280
Db	2221	GCAGCAATACCATTTGACCTCCGATGTCCCCCGGACCTTTACATTCATTGATGACAGTTG	2280
QY	2281	GAATTAACGGGTGAAGTGTCTGAGACCCCTCAGCAGGAGGCGAGTTTGAAGTGTGACGTTTG	2340
Db	2281	GAATTAACGGGTGAAGTGTCTGAGACCCCTCAGCAGGAGGCGAGTTTGAAGTGTGACGTTTG	2340
QY	2341	ACATGTGATCTGACCTCGGAGTGTGTACTCTCCGCCATGTGAGAGAGCTGAACCCAGAGCT	2400
Db	2341	ACATGTGATCTGACCTCGGAGTGTGTACTCTCCGCCATGTGAGAGAGCTGAACCCAGAGCT	2400
QY	2401	GCAGGAGCTGACTTGTGAGACACCTGCCCGGTGCTCCACCCTTAACGAGCCGAAACCCATA	2460
Db	2401	GCAGGAGCTGACTTGTGAGACACCTGCCCGGTGCTCCACCCTTAACGAGCCGAAACCCATA	2460
QY	2461	TGCTGTGAACCTCCTAACTTTGTGTTCAGATTTTTTTTTTAAATTCCTACTTGTCT	2520
Db	2461	TGCTGTGAACCTCCTAACTTTGTGTTCAGATTTTTTTTTTAAATTCCTACTTGTCT	2520
QY	2521	ATCTTTGGGCAATGTGGGCACTTTTAAAAAGAGAAATGAGTGAAGTGGGTGATTAAC	2580
Db	2521	ATCTTTGGGCAATGTGGGCACTTTTAAAAAGAGAAATGAGTGAAGTGGGTGATTAAC	2580
QY	2581	TGTTATGTAAAGAGAGAGACCTCTGATGTGGGATGGGGCTGAGAGCAGAGGGAGGC	2640
Db	2581	TGTTATGTAAAGAGAGAGACCTCTGATGTGGGATGGGGCTGAGAGCAGAGGGAGGC	2640
QY	2641	AAAGGACACGCTCTGCTCTGCTCCCGCGCCGCTCCTTTTCACAGAGCTGGGGGTTGG	2700
Db	2641	AAAGGAGACACGCTCTGCTCTGCTCCCGCGCTCCTTTTCACAGAGCTGGGGGTTGG	2700
QY	2701	TTGTTAGACAAGTGCCTCTGTGTGGCCATGTGCTACCTGTGGCCGACCTCTGTGAGCTGAT	2760
Db	2701	TTGTTAGACAAGTGCCTCTGTGTGGCCATGTGCTACCTGTGTGGCCGACCTGTGAGCTGAT	2760
QY	2761	ACCCCATTTCTGGGAAACCTCGGTCTGACACTTTCGAACCTTGTAATATCCACATGAAAGC	2820
Db	2761	ACCCCATTTCTGGGAAACCTCGGTCTGACACTTTCGAACCTTGTAATATATCCACATGAAAGC	2820
QY	2821	TAGGACTAAGCCCGAGAGGTTCTCTCTTTAAATTTAAAAAATTTTTAAAAAATTTTTAAAAA	2889
Db	2821	TAGGACTAAGCCCGAGAGGTTCTCTCTTTAAATTTAAAAAATTTTTAAAAAATTTTTAAAAA	2889
RESULT	3		
LOCUS	AR121061	2869 bp	DNA
DEFINITION	Sequence 82 from patent US 6159694.		
ACCESSION	AR121061		
KEYWORDS	.AR121061.1 GI:14104637		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2869)		
AUTHORS	Kariras,J.G.		
TITLE	Antisense modulation of stat3 expression		
JOURNAL	Patent: US 6159694-A 82 12-DEC-2000;		
FEATURES	Location/Qualifiers		
	source	1..2869	
		/organism="unknown"	
BASE COUNT	755 a	743 c	772 g
ORIGIN		599 t	

	Query Match	99.8%:	Score 2864.2:	DB 6:	Length 2869:	
	Best Local Similarity	99.9%:	Pred. No.0:	Mismatches	3:	Indels
	Matches 2866:	Conservative	0:			Gaps
OY	1	GCCTGACAGCAGCAGCCGGCCATGTGGGGCTTCAGCCCGAGACAGTGTGAGACCTCTGACT	60			
Db	1	GCCTGACAGCAGCAGCCGGCCATGTGGGGCTTCAGCCCGAGACAGTGTGAGACCTCTGACT	60			
OY	61	GCAGCAGATGGCTCAGTGTGAACCCAGCTGCACGACCTGTGACAGCTATCCGTGAAGCAGC	120			
Db	61	GCAGCAGATGGCTCAGTGTGAACCCAGCTGCACGACCTGTGACAGCTATCCGTGAAGCAGC	120			
OY	121	TGCACACAGCTGTACAGCGCACGACGTTCCCATGTGAGACTGTGGCAGTTCCTGGCACCTTGA	180			
Db	121	TGCACACAGCTGTACAGCGCACGACGTTCCCATGTGAGACTGTGGCAGTTCCTGGCACCTTGA	180			
OY	181	TTGACAGTCACAGCTGTGGCATATTCGACGCCAGCAAGATGCACATGCCAGTGTGTTTC	240			
Db	181	TTGACAGTCACAGCTGTGGCATATTCGACGCCAGCAAGATGCACATGCCAGTGTGTTTC	240			
OY	241	ATTAATCTCTTGGGTAAATTTGACCGACATATAGCCGATTCCTGCAGAGTCCCAATGTCC	300			
Db	241	ATTAATCTCTTGGGTAAATTTGACCGACATATAGCCGATTCCTGCAGAGTCCCAATGTCC	300			
OY	301	TCTATCAGCACACCTTGCAGAAATCCAGCACTTTCGACAGCAGTATCTTGAAGAC	360			
Db	301	TCTATCAGCACACCTTGCAGAAATCCAGCACTTTCGACAGCAGTATCTTGAAGAAC	360			
OY	361	CAATGAAATTTCCCGGATCTGTGCCGATGTGCTGTGGGAAAGTTCCTCTCTCCAGA	420			
Db	361	CAATGAAATTTCCCGGATCTGTGCCGATGTGCTGTGGGAAAGTTCCTCTCTCCAGA	420			
OY	421	CGGCAAGCCAGCGCAGCCAGCAAGGGGGCCAGCCCAACCAACCAAGCCGCTGATGA	480			
Db	421	CGGCAAGCCAGCGCAGCCAGCAAGGGGGCCAGCCCAACCAACCAAGCCGCTGATGA	480			
OY	481	CAGAGAGCAGAGATGTTGGAGCAGCATCTTCAGATGTCCGGAAAGCAGTGTGACAGATC	540			
Db	481	CAGAGAGCAGAGATGTTGGAGCAGCATCTTCAGATGTCCGGAAAGCAGTGTGACAGATC	540			
OY	541	TAGAACAGAAAAATGAAGTGTGTGAGAACCTCCAGGACGACTTGTATTTCAACTACAAA	600			
Db	541	TAGAACAGAAAAATGAAGTGTGTGAGAACCTCCAGGACGACTTGTATTTCAACTACAAA	600			
OY	601	CCCTCAAGGCCAAGGAGACATGACAGATCTGAATGGAACCAACCGTGTGACACGAC	660			
Db	601	CCCTCAAGGCCAAGGAGACATGACAGATCTGAATGGAACCAACCGTGTGACACGAC	660			
OY	661	AGAAGATGAGCAGCTGTGAAGACAGATCTCACAGCCCTTGACACGATGCGAGAACGATTG	720			
Db	661	AGAAGATGAGCAGCTGTGAAGACAGATCTCACAGCCCTTGACACGATGCGAGAACGATTG	720			
OY	721	TGAGTACAGCTGGCGGGGCTTCTCAGCAATGAGTACGTGTGCAAGAACATGACTGATG	780			
Db	721	TGAGTACAGCTGGCGGGGCTTCTCAGCAATGAGTACGTGTGCAAGAACATGACTGATG	780			
OY	781	AAGAGCTGCTGACTGTGAAGAGCGGCCACAGATCTCGTGTGATCGGAGGCCCTCCCAACA	840			
Db	781	AAGAGCTGCTGACTGTGAAGAGCGGCCACAGATCTCGTGTGATCGGAGGCCCTCCCAACA	840			
OY	841	TCTGCTGACCGCTCTGGAATACTGATTAATTCATTAGCAGATTTCAACTTCAGACCC	900			
Db	841	TCTGCTGACCGCTCTGGAATACTGATTAATTCATTAGCAGATTTCAACTTCAGACCC	900			
OY	901	GCCAAACAATTTAAGAACTGTGAGAGCTCAGAGAAAGTGTCTTCAAGGGGAGCCTTA	960			
Db	901	GCCAAACAATTTAAGAACTGTGAGAGCTCAGAGAAAGTGTCTTCAAGGGGAGCCTTA	960			
OY	961	TGCTGACAGACCGGCCCATGTGTGAGAGAGAGATCTGTGAGCTGTTTCAGAAACTTAATGA	1020			
Db	961	TGCTGACAGACCGGCCCATGTGTGAGAGAGAGATCTGTGAGCTGTTTCAGAAACTTAATGA	1020			

QY	1021	AGATGCTCTTCGTGTCGACCGGACCGCTCGATGCGCATGACCGCGACCGGCTTAG	1080
QY	1021	AGATGCTCTTCGTGTCGACCGGACCGCTCGATGCGCATGACCGCGACCGGCTTAG	1080
Db	1021	AGATGCTCTTCGTGTCGACCGGACCGCTCGATGCGCATGACCGCGACCGGCTTAG	1080
QY	1081	TCATCAAGACATGCGTGTCCAGTTTACCACGAAATGCAAGTTTGGTCAAAATTTCTCGAGT	1140
Db	1081	TCATCAAGACATGCGTGTCCAGTTTACCACGAAATGCAAGTTTGGTCAAAATTTCTCGAGT	1140
QY	1141	TGAAATTTATCAGCTTAAATTTAAAGTGTCATTTGATATAAACAATCTGGGATGTGTGTCGC	1200
Db	1141	TGAAATTTATCAGCTTAAATTTAAAGTGTCATTTGATATAAACAATCTGGGATGTGTGTCGC	1200
QY	1201	TCAGAGGCTTCGCGAAATTTAAACATTTCTGGGACGCAACCAAAAGATGATGAACATGACAG	1260
Db	1201	TCAGAGGCTTCGCGAAATTTAAACATTTCTGGGACGCAACCAAAAGATGATGAACATGACAG	1260
QY	1261	AGCTCTCAAGGCGAGCTGTGTGTCAGAGTTTCAAGCACTGACCTTAGGAGGACAGACAT	1320
Db	1261	AGCTCTCAAGGCGAGCTGTGTGTCAGAGTTTCAAGCACTGACCTTAGGAGGACAGACAT	1320
QY	1321	GTGGGAATGGAGGCGCGGCCAATTTGTGATCCCTCTTGATTCGATGTAGAGAGCTGCACC	1380
Db	1321	GTGGGAATGGAGGCGCGGCCAATTTGTGATCCCTCTTGATTCGATGTAGAGAGCTGCACC	1380
QY	1381	TGATCACCCTTCGAGACTGAGGTGTACCAACAGGCTCAAGATTGACCTAGAGACCCACT	1440
Db	1381	TGATCACCCTTCGAGACTGAGGTGTACCAACAGGCTCAAGATTGACCTAGAGACCCACT	1440
QY	1441	CCTTGGCAGTTGTGTGTATCTCCACACATCTGTAGATGGCAAAATGTTGGGATCAATCC	1500
Db	1441	CCTTGGCAGTTGTGTGTATCTCCACACATCTGTAGATGGCAAAATGTTGGGATCAATCC	1500
QY	1501	TGTGTATTAACATGCTCTGACCAATTAACCCCAAGAGCTGAACCTTCTCACTAAGCGCGAA	1560
Db	1501	TGTGTATTAACATGCTCTGACCAATTAACCCCAAGAGCTGAACCTTCTCACTAAGCGCGAA	1560
QY	1561	TTTGAACCTGGGACCAAGTGGCGAGGTGCTCAGCTGTGCAAGTTCTGTGTCACACCAACGC	1620
Db	1561	TTTGAACCTGGGACCAAGTGGCGAGGTGCTCAGCTGTGCAAGTTCTGTGTCACACCAACGC	1620
QY	1621	GAGGCTCAGCATGAGACAGCTGACACAGCTGCGCTGAGAAGTCTCTAAGGCGCTGTGTGTA	1680
Db	1621	GAGGCTCAGCATGAGACAGCTGACACAGCTGCGCTGAGAAGTCTCTAAGGCGCTGTGTGTA	1680
QY	1681	ACTACTCAGGCTGTACATCATCATGTGGCTAAATTTCTGCAAAAGAAACATGTGCTGGCAAG	1740
Db	1681	ACTACTCAGGCTGTACATCATCATGTGGCTAAATTTCTGCAAAAGAAACATGTGCTGGCAAG	1740
QY	1741	GCTTCTCCTTGTGGGTGGCTGTACAAATTAATCATCGACTTGTGAAAAGTAAATCTTGG	1800
Db	1741	GCTTCTCCTTGTGGGTGGCTGTACAAATTAATCATCGACTTGTGAAAAGTAAATCTTGG	1800
QY	1801	CCCTTTGGAATGAAGGGTATCATCTAGGCTTTCATCAGCAAGAAGCGGAGCGGCGCATCC	1860
Db	1801	CCCTTTGGAATGAAGGGTATCATCTAGGCTTTCATCAGCAAGAAGCGGAGCGGCGCATCC	1860
QY	1861	TAAAGCAAAACCCCCCGGACCTTCTACTGCGCTTCACGCGAGACAGCAAAAGAAAGAG	1920
Db	1861	TAAAGCAAAACCCCCCGGACCTTCTACTGCGCTTCACGCGAGACAGCAAAAGAAAGAG	1920
QY	1921	GGGTACACTTTCATTTGGGTGGAAAGAGACATCAGTGTGGCAAGACCCAGATCCAGTCTGTAG	1980
Db	1921	GGGTACACTTTCATTTGGGTGGAAAGAGACATCAGTGTGGCAAGACCCAGATCCAGTCTGTAG	1980
QY	1981	AGCCATACACCAAGCAGCAGTGAACAACATGTCTATTGTGAAATCATCATGTGGCTATATA	2040
Db	1981	AGCCATACACCAAGCAGCAGTGAACAACATGTCTATTGTGAAATCATCATGTGGCTATATA	2040
QY	2041	AGATCATGTGATGCGACCAACATCTCTGGTGTCTGCACCTTGTCTACCTCTAACCCGCACATTC	2100
Db	2041	AGATCATGTGATGCGACCAACATCTCTGGTGTCTGCACCTTGTCTACCTCTAACCCGCACATTC	2100
QY	2101	CCAAAGAGAGGAGCAATTTTGAAAGTACTGTAGGCCCGAGACCGACGAGACACCCCGAAACCG	2160

Db	2101	CCAGAGGAGGCATTTCGGAAAGTACTGTAGGCCCGCAGACGACGACACCCTGAAGCCG	2160
OY	2161	ACCCAGGATATGCTGCCCGCTACCTGAAAGACCAAGTTGATCTGTGTGACACCAAGACT	2220
Db	2161	ACCCAGGATATGCTGCCCGCTACTTAAGAAGCACCAAGTTCATCTGTGTACACCAAGACT	2220
OY	2221	GCAGCATATACCATTTGACCTGCCGATATGCTCCCCCGCAGCTTGATATCATGATGACGTTTG	2280
Db	2221	GCACCAATATACCATTTGACCTGCCGATATGCTCCCCCGCAGCTTGATATCATGATGACGTTTG	2280
OY	2281	GAAATAACGGTGAAGTNGCTGAGACCCCTCAGCAGAGAGGAGCTTTTAGTGCCTCAGCTTTG	2340
Db	2281	GAAATAACGGTGAAGTNGCTGAGACCCCTCAGCAGAGAGGAGCTTTTAGTGCCTCAGCTTTG	2340
OY	2341	ACATGATCTGACCTGGGAGTGTGCTACTCTCCCCTATGTGAGAGACTGAAACCGAAGCT	2400
Db	2341	ACATGATCTGACCTGGGAGTGTGCTACTCTCCCCTATGTGAGAGACTGAAACCGAAGCT	2400
OY	2401	GCAGAGACGTGACTTGAGACACCTGCCCGCTCCAGCACCCTCAGACAGCCGAAACCCATA	2460
Db	2401	GCAGAGACGTGACTTGAGACACCTGCCCGCTCCAGCACCCTCAGACAGCCGAAACCCATA	2460
OY	2461	TGCTGTCAAACCTCCTAACCTTTGTGTTCCAGATTTTTTTTTTAAATTTCTTACTTCTCT	2520
Db	2461	TGCTGTCAAACCTCCTAACCTTTGTGTTCCAGATTTTTTTTTTAAATTTCTTACTTCTCT	2520
OY	2521	ATCTTTGGGCAATCTGGGCACTTTTAAAAAGAGAAATAGCTGAGTGGGTGATTAAC	2580
Db	2521	ATCTTTGGGCAATCTGGGCACTTTTAAAAAGAGAAATAGCTGAGTGGGTGATTAAC	2580
OY	2581	TGTTATGTAAAGAGAGAGACCTCTGATGTCGGGATGGGGCTGAGAGCAAGAGGAGGC	2640
Db	2581	TGTTATGTAAAGAGAGAGACCTCTGATGTCGGGATGGGGCTGAGAGCAAGAGGAGGC	2640
OY	2641	AAAAGGGAACAACCTCCGTCTGCTCCGCCGCTGCCCTCTTTTACAGAGCTGGGGGGTTGG	2700
Db	2641	AAAAGGGAACAACCTCCGTCTGCTCCGCCGCTGCCCTCTTTTACAGAGCTGGGGGGTTGG	2700
OY	2701	TTGTTAGACAAGTGCCTCCTGATGCCATGAGCTTGTGGCCCACTCTGTGAGCTGAT	2760
Db	2701	TTGTTAGACAAGTGCCTCCTGATGCCATGAGCTTGTGGCCCACTCTGTGAGCTGAT	2760
OY	2761	ACCCATTTCTGGGAACCTCTGGCTCTGACCTTTCATCTTGCATATGCCATGAAAGC	2820
Db	2761	ACCCATTTCTGGGAACCTCTGGCTCTGACCTTTCATCTTGCATATGCCATGAAAGC	2820
OY	2821	TAGAGCTAAGCCCGCAGAGTTCCTCTTTAAATTAAAAAAAAAAAAAAAAAAAAAA	2869
Db	2821	TAGAGCTAAGCCCGCAGAGTTCCTCTTTAAATTAAAAAAAAAAAAAAAAAAAAAA	2869
RESULT 4			
MMU06922			
LOCUS	2869 bp	mRNA	01-JUL-1994
DEFINITION	Mus musculus signal transducer and activator of transcription		
VERSION	U06922.1		
KEYWORDS	GI:458705		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
TITLE	Zhong,Z., Wen,Z., and Darnell,J.E.		
JOURNAL	Stats: a STAT family member activated by tyrosine phosphorylation		
MEDLINE	In response to epidermal growth factor and interleukin-6		
REFERENCE	Science 264, 95-98 (1994)		
AUTHORS	94188718		
TITLE	2 (bases 1 to 2869)		
	Zhong,Z., Wen,Z., and Darnell,J.		
	Stats and Stats: members of the family of signal transducers and		

JOURNAL	activators of transcription
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 91, 4806-4810 (1994)
REFERENCE	94255416
AUTHORS	3 (bases 1 to 2869)
TITLE	Zhong, Z.
JOURNAL	Direct Submission
	Submitted (18-FEB-1994) Z. Zhong, The Rockefeller University, The
	Molecular Cell Biology Laboratory, 1230 York Avenue, New York, NY
	10021, USA
FEATURES	Location/Qualifiers
source	1..2869
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/sex="female"
	/tissue_type="thymus"
	/clone_id="Mouse thymus cDNA library (Stratagene)"
	/dev_stage="6-8 weeks"
gene	69..2381
	/gene="Stat3"
CDS	69..2381
	/gene="Stat3"
	/function="DNA binding protein induced by EGF and IL-6"
	/citation={1}
	/citation={2}
	/codon_start=1
	/evidence=experimental
	/product="Stat3"
	/protein_id="AA19452.1"
	/db_xref="GI:458706"
	/translation="MAQNNQLOQDPTVYIKIOLHYSDETPMELROFLAPMISQDA YAAKSEHATLVFHNHLEIDQOTSRLEDSNVLTQHNLRIRKIFLOSRILEKPMELA RIVAKLWEESRLQTAATAAQGGQAHPTAAVATVTRKQMLEOHLDVRRKRODLEO KIMVVENIADQDFPNYKTLKSGQMODLNGNSQVTRKQMOLEOMLTALDQMRKSTV SELQGLISAMEYVOKLTDEELADMRKROQJACIGGPNICLRLKLEMLITSLAESQLO TRQOIKLKEELOOKVSYKGDPIYOHPRMLEERIVLEPRNLKMSAFVEVRQPMHDP RPIYKTVGTFQTVRLLVYKREPNLTOLKIKVCIDKSGVAAALRGSRRNIICTNKR VNNESNNGSLSAEFKLTILREORCGNGRAGNDASLIYTELHLITPEVEYHGLL KIDNETHSLPVAIVISNICOHPNMAASLIYMLTNRNKNFTKPIGTHQVDAEVL SMQSFSTTKRGSLIEQLTTAEKLLGPNVTSQGIITWAFCKENMGKSPFVWLD NITLDVKKYKILALNEGIMGFISKERRALLSTRPGPTFLRSESSKSGGVFTWV EKDLSGKQIOISVEPYTKQOLNNMSFAEIIINGYKIMDATNLISPLVLYDPIPKERA FGKCRPSEHPEADPGSAAPVYKTKRFICVPTPTCSNTIDLPMSPRTLSLMQGNNN GEGAPRSAGGQFESLTFPMDLTSCATSPM"
BASE COUNT	755 a 743 c 772 g 599 t
ORIGIN	
Query Match	99.8%; Score 2864.2; DB 10; Length 2869;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 2866; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
QY	1 GCCGCGACCCAGCGAGCGCGGCGGCTCAGCCCGGAGCAGCTCGAGACCCCTGACT 60
DB	1 GCCGCGACCCAGCGAGCGCGGCGGCTCAGCCCGGAGCAGCTCGAGACCCCTGACT 60
QY	61 GGAGCAGATGGCTCAGTGTGAAGCAGCTCAGCAGCTGGAGCAGCAGCTAAGTGAAGCAGC 120
DB	61 GGAGCAGATGGCTCAGTGTGAAGCAGCTCAGCAGCTGGAGCAGCAGCTAAGTGAAGCAGC 120
QY	121 TGCACAGCTGTACAGCGCAGCGTTCCCATGGAGCTGGCGCAGTTCCTGGACACCTTGA 180
DB	121 TGCACAGCTGTACAGCGCAGCGTTCCCATGGAGCTGGCGCAGTTCCTGGACACCTTGA 180
QY	181 TTGAGAGTGAAGACGTGGGATATGACAGCCAGCAAGAGATGCACATCCCATTTGGTGTTC 240
DB	181 TTGAGAGTGAAGACGTGGGATATGACAGCCAGCAAGAGATGCACATCCCATTTGGTGTTC 240
QY	241 ATAATCTCTTGGGTGAATAATTGACACAGCAATATAGCGATTCCTGCAAGAGTCCAAATGCC 300
DB	241 ATAATCTCTTGGGTGAATAATTGACACAGCAATATAGCGATTCCTGCAAGAGTCCAAATGCC 300
QY	301 TCTATCAGACAACCTTCGAAGAATCAACAGATTTCTGACAGCAGGTATCTTGAGAAAC 360
DB	301 TCTATCAGACAACCTTCGAAGAATCAACAGATTTCTGACAGCAGGTATCTTGAGAAAC 360

Db 301 TCATCAGCAACCTTCGAAAGATCAAGCACTTCTGCGAGAGCAGTATCTTAGAACC 360
Oy 361 CAATGAAATTTGCCCGGATCGTGGCCCCGATGCTGTGGGAAGTCTCGCCTCCGCCAGA 420
Db 361 CAATGAAATTTGCCCGGATCGTGGCCCCGATGCTGTGGGAAGTCTCGCCTCCGCCAGA 420
Oy 421 CGCGAGCCAGCGGACGCCAGAGGGGGCCAGGCCAACACCCCAAGCCGCCCTACTGA 480
Db 421 CGCGAGCCAGCGGACGCCAGAGGGGGCCAGGCCAACACCCCAAGCCGCCCTACTGA 480
Oy 481 CAGAGAAGCAGCAGATGTGGAGCAGCATCTTCAGATGTCCGAGACGAGTGCAGATC 540
Db 481 CAGAGAAGCAGCAGATGTGGAGCAGCATCTTCAGATGTCCGAGAGGAGATGCAGATC 540
Oy 541 TAGAAGCAAAAAATGAAGTGTGGAGAACTCCAGACGACTTTGATTTCACTACAAA 600
Db 541 TAGAAGCAAAAAATGAAGTGTGGAGAACTCCAGACGACTTTGATTTCACTACAAA 600
Oy 601 CCCCAAGAGCCAGAGACATGAGATCGAGATCGAATGGAACACGAGTGTGACCAAGC 660
Db 601 CCCCAAGAGCCAGAGACATGAGATCGAGATCGAATGGAACACGAGTGTGACCAAGC 660
Oy 661 AGAAGATGCAGACACTGGAACAGATGCTCACAGCCCTGGACAGATGGCGAAGCAATG 720
Db 661 AGAAGATGCAGACACTGGAACAGATGCTCACAGCCCTGGACAGATGGCGAAGCAATG 720
Oy 721 TGAATGAGCTGGCGGGCTCTTTTCAGCAATGAGTACGTGCAGAAAGCACTGACTGATG 780
Db 721 TGAATGAGCTGGCGGGCTCTTTTCAGCAATGAGTACGTGCAGAAAGCACTGACTGATG 780
Oy 781 AAGAGCTGGCTGACTGGAAGAGGGCGGCGAGATCGCTGATGGAAGGCCCTCCACA 840
Db 781 AAGAGCTGGCTGACTGGAAGAGGGCGGCGAGATCGCTGATGGAAGGCCCTCCACA 840
Oy 841 TCTGCTGAGACCGTCTGGAAGAACTGGATACCTTATTAGCAATCTCAACTTACAGACC 900
Db 841 TCTGCTGAGACCGTCTGGAAGAACTGGATACCTTATTAGCAATCTCAACTTACAGACC 900
Oy 901 GCCAACAATTAAGAAACTGGAGAGCTGCAGCAAAAGTGTCTCTACAAAGGCCACCTA 960
Db 901 GCCAACAATTAAGAAACTGGAGAGCTGCAGCAAAAGTGTCTCTACAAAGGCCACCTA 960
Oy 961 TCGTCAGACACGGGCCCATGCTGAGAGAGAGATCGTGAGACTTTCAGAACTTAATGA 1020
Db 961 TCGTCAGACACGGGCCCATGCTGAGAGAGAGATCGTGAGACTTTCAGAACTTAATGA 1020
Oy 1021 AGAGTGCCTTGTGTGGAGGGGCGAGCCCTGATGCCATGCACCCGAGCCGCCCTTAG 1080
Db 1021 AGAGTGCCTTGTGTGGAGGGGCGAGCCCTGATGCCATGCACCCGAGCCGCCCTTAG 1080
Oy 1081 TCATCAACAGCTGTGTCCAGTTTACACAGAAAGTCAGGTCTGTGTCAAATTTCTGAGT 1140
Db 1081 TCATCAACAGCTGTGTCCAGTTTACACAGAAAGTCAGGTCTGTGTCAAATTTCTGAGT 1140
Oy 1141 TGAATTAACACTTAAATTAAGTGTGATTAAGTAAAGACTGTGGGATGTGTGCCC 1200
Db 1141 TGAATTAACACTTAAATTAAGTGTGATTAAGTAAAGACTGTGGGATGTGTGCCC 1200
Oy 1201 TCAGAGGGTCTCGGAAATTTAACTTCTGGGACAGAACACAAAAGTGAATGAGGAGG 1260
Db 1201 TCAGAGGGTCTCGGAAATTTAACTTCTGGGACAGAACACAAAAGTGAATGAGGAGG 1260
Oy 1261 AGTCAACACAGCGAGCTGTCTGAGAGTTCAAGACACTTACCTTAGAGAGCAGAT 1320
Db 1261 AGTCAACACAGCGAGCTGTCTGAGAGTTCAAGACACTTACCTTAGAGAGCAGAT 1320
Oy 1321 GTGGGAATGAGGCGCGTGTCCATTTGTATGCTCTCTGATGTGATGAGGAGCGTGCACC 1380
Db 1321 GTGGGAATGAGGCGCGTGTCCATTTGTATGCTCTCTGATGTGATGAGGAGCGTGCACC 1380
Oy 1381 TGATCACTTGAAGCTGAGTGTACACCAAGGCTCAAGATTGACCTAGAGACCCACT 1440
Db 1381 TGATCACTTGAAGCTGAGTGTACACCAAGGCTCAAGATTGACCTAGAGACCCACT 1440

Oy 1441 CTTTCCAGTTGTGTATCTCCAACTGTGTGAGATGCCAAATGCTTGGGATCAATCC 1500
Db 1441 CTTTCCAGTTGTGTATCTCCAACTGTGTGAGATGCCAAATGCTTGGGATCAATCC 1500
Oy 1501 TGTGTAATACATGCTGACCAATTAACCCCAAGAGCTGAACCTTCTACTAAGCCGCAA 1560
Db 1501 TGTGTAATACATGCTGACCAATTAACCCCAAGAGCTGAACCTTCTACTAAGCCGCAA 1560
Oy 1561 TTTGAACCTGGAGCAAGTGGCGGAGTGTCTACCTGGCAGTCTCTCCACACCAAC 1620
Db 1561 TTTGAACCTGGAGCAAGTGGCGGAGTGTCTACCTGGCAGTCTCTCCACACCAAC 1620
Oy 1621 GAGGCTGAGCATGACAGCTGACCAAGCTGGCTGGAAGCTCTAGGCGCTGTGTGA 1680
Db 1621 GAGGCTGAGCATGACAGCTGACCAAGCTGGCTGGAAGCTCTAGGCGCTGTGTGA 1680
Oy 1681 ACTACTAGGCTGTACATCATGAGGCTAAATCTGCAAGAAACATGCTGGCAAG 1740
Db 1681 ACTACTAGGCTGTACATCATGAGGCTAAATCTGCAAGAAACATGCTGGCAAG 1740
Oy 1741 GCTTCTCTTGTGGTCTGGCTAGACAAATATCATGACCTTGTGAAAAAGTATCTTGG 1800
Db 1741 GCTTCTCTTGTGGTCTGGCTAGACAAATATCATGACCTTGTGAAAAAGTATCTTGG 1800
Oy 1801 CCCTTTGAATGAAGGCTACATCATGAGGCTTATCAGCAAGAGCGGGAGCGGCCATTC 1860
Db 1801 CCCTTTGAATGAAGGCTACATCATGAGGCTTATCAGCAAGAGCGGGAGCGGCCATTC 1860
Oy 1861 TAAACCAAAAGCCCCCGGGACCTTCTACTGCGCTTACGAGAGACGCAAAAGAGAG 1920
Db 1861 TAAACCAAAAGCCCCCGGGACCTTCTACTGCGCTTACGAGAGACGCAAAAGAGAG 1920
Oy 1921 GGGTCACTTACTTGGTGGTGAAGAAAGACATCATGAGCAACCAAGTCCAGTCTGAG 1980
Db 1921 GGGTCACTTACTTGGTGGTGAAGAAAGACATCATGAGCAACCAAGTCCAGTCTGAG 1980
Oy 1981 AGCCATACCAACAGCAGCTGAAACAATGTATTTGCTGAATCATCATGAGGCTATA 2040
Db 1981 AGCCATACCAACAGCAGCTGAAACAATGTATTTGCTGAATCATCATGAGGCTATA 2040
Oy 2041 AGATCATGATGCGACCAACATCTGTGTCTCCACTTGTACTTACCTTACCCGCAATTC 2100
Db 2041 AGATCATGATGCGACCAACATCTGTGTCTCCACTTGTACTTACCTTACCCGCAATTC 2100
Oy 2101 CCAAGAGAGGAGCAATTGGAAGTACTGTAGGCCGAGAGCCAGAGCACCAGGAGCG 2160
Db 2101 CCAAGAGAGGAGCAATTGGAAGTACTGTAGGCCGAGAGCCAGAGCACCAGGAGCG 2160
Oy 2161 ACCAGTAGTGTGCCCCGCTACCTGAAGACCAAGTTCATGTGTGACACCAAGACT 2220
Db 2161 ACCAGTAGTGTGCCCCGCTACCTGAAGACCAAGTTCATGTGTGACACCAAGACT 2220
Oy 2221 GCAGCAATACCATTTGACCTGCGATGTCGCCGACCTTGTGATTCATTTGATGAGGAT 2280
Db 2221 GCAGCAATACCATTTGACCTGCGATGTCGCCGACCTTGTGATTCATTTGATGAGGAT 2280
Oy 2281 GAAATTAACGTTGAAGTGTGAGCTTACCTGAGGAGGAGGAGGAGGAGTGTGAGTGCACGTTTG 2340
Db 2281 GAAATTAACGTTGAAGTGTGAGCTTACCTGAGGAGGAGGAGGAGGAGTGTGAGTGCACGTTTG 2340
Oy 2341 ACATGATCTGACCTCGGAGTGTGCTACCTGCTCCCATGTGAGAGCTGAACCGAAGCT 2400
Db 2341 ACATGATCTGACCTCGGAGTGTGCTACCTGCTCCCATGTGAGAGCTGAACCGAAGCT 2400
Oy 2401 GCAGAGACTGACTTGAAGACCTGCCCCGCTGCTCAACCCCTAAGCAGCCGACCCGATA 2460
Db 2401 GCAGAGACTGACTTGAAGACCTGCCCCGCTGCTCAACCCCTAAGCAGCCGACCCGATA 2460
Oy 2461 TCGTCTGAACCTCCTACTTGTGTGCTTCAGATTTTCTTATTTCTACTTCTGCT 2520
Db 2461 TCGTCTGAACCTCCTACTTGTGTGCTTCAGATTTTCTTATTTCTACTTCTGCT 2520

Oy	2521	ATCTTTGGCAATCTGGGCACTTTTAAAAAGAGAAATGAGTGAGTGGGTATTAAC	2580
Db	2521	ATCTTTGGGCAATCTGGGCCTTTTAAAAGAGAAATGAGTGAGTGGGTATTAAC	2580
Oy	2581	TGTTATTAAGAAGAGAGACTCTTAGTCGTGGGATGGGGCTGAGACAGAAAGGAGGC	2640
Db	2581	TGTTATGAAGAAGAGAGACTCTTAGTCGTGGGATGGGGCTGAGACAGAAAGGAGGC	2640
Oy	2641	AAAGGGACAACCCTCGTCCTGCCGGCCGCTCCTTTTACACAGCTGGGGGTGG	2700
Db	2641	AAAGGGACAACCCTCGTCCTGCCGGCCGCTCCTTTTACACAGCTGGGGGTGG	2700
Oy	2701	TTGTAGACAAGTCCCTCTGTCGTGTCGCCATGCGTACCTGTTGCCCACTCTGTGAGCTGAT	2760
Db	2701	TTGTAGACAAGTCCCTCTGTCGTGTCGCCATGCGTACCTGTTGCCCACTCTGTGAGCTGAT	2760
Oy	2761	AACCATTCCTGGGAACTCGTGGGCTGTCGATTTTCACACCTTGCCTAATATCCACATAGGAAC	2820
Db	2761	AACCATTCCTGGGAACTCGTGGGCTGTCGATTTTCACACCTTGCCTAATATCCACATAGGAAC	2820
Oy	2821	TAGAGCTAAGCCACGAGAGTTCTCTTTAAATTAAAAAAAAAAAAAAAAAAAA	2869
Db	2821	TAGAGCTAAGCCACGAGAGTTCTCTTTAAATTAAAAAAAAAAAAAAAAAAAA	2869

RESULT 5
 ARI153599
 LOCUS ARI153599 2869 bp DNA PAT 08-AUG-2001
 DEFINITION Sequence 9 from patent US 6235873.
 ACCESSION ARI153599
 VERSION ARI153599.1 GI:15121131
 KEYWORDS
 SOURCE
 ORGANISM
 UNKNOWN.
 UNCLASSIFIED.
 REFERENCE 1 (bases 1 to 2869)
 AUTHORS Bromberg,J.F., Wrzeszczynska,M.H., Zhao,Y. and Darnell,J.E. Jr
 TITLE Constitutively active transcription factors and their uses for
 identifying modulators of activity including dysproliferative
 cellular changes
 JOURNAL Patent: US 6235873-A 9 22-MAY-2001;
 FEATURES
 source location/qualifiers
 1..2869
 /organism="unknown"
 BASE COUNT 753 a 742 c 772 g 602 t
 ORIGIN

Query Match	99.7%	Score 2861;	DB 6;	Length 2869;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 2864; Conservative	0;	Mismatches	5;	Indels 0; Gaps 0;

QY	1	GGCGGG-CCGAGCCAGCGGGCCAGTGGGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT	60
	11	GGCGGGGACCGCAGCAGCGGGGCGAGTGGGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT	60
Db	1	GGCGGGGACCGCAGCAGCGGGGCGAGTGGGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT	60
QY	61	GCAGCAGATGGCTAGTGGAAACAGGTGGAGCAGTGGACACAGCTACTGTAAGACAGC	120
Db	61	GCAGCAGATGGCTAGTGGAAACAGGTGGAGCAGTGGACAGCTGGACACAGCTACTGTAAGACAGC	120
QY	121	TGCACCACTGTACAGGCACACGTTCCCATGAGCTGGGGCAAGTTCTTGACACTTGGGA	180
Db	121	TGCACCACTGTGTACAGGCACACGTTCCCATGAGCTGGGGCAAGTTCTTGACACTTGGGA	180
QY	181	TTCGAGAGTCAACACTGGGGCATATGCAGGCACAGCAAGAGTACATATGCACAGTTGGTTTC	240
Db	181	TTCGAGAGTCAACACTGGGGCATATGCAGGCACAGCAAGAGTACATATGCACAGTTGGTTTC	240
QY	241	ATPATCTCTGGGGTGAATTGACCAGCAGATATAGCCGATTTCTTGCAAGAGTCCATGTCC	300
Db	241	ATPATCTCTGGGGTGAATTGACCAGCAGATATAGCCGATTTCTTGCAAGAGTCCATGTCC	300
QY	301	TCTATTCAGACACCTTGGAAAGTAACAGCGTTTCTGCAGACAGATATCTTGAAGAC	360

Db	301	TCATATCGACAAACCTTCGAAAGATCAACGACTTTCGTGAGAGCAGATATCTTGAAAGC	360
Qy	361	CAATGGAAATTGCCGGGATCTGTGGCCCCGATGCCCTGTGGAAAGTCTGGCTCTCTCCAGA	420
Db	361	CAATGGAAATTGCCGGGATCTGTGGCCCCGATGCCCTGTGGAAAGTCTGGCTCTCTCCAGA	420
Qy	421	CGGAGACGACGAGCCACGACCAAGGGGGCCAGGCCCAACCCCAACAGCCGCTACTGA	480
Db	421	CGGAGACGACGAGCCACGACCAAGGGGGCCAGGCCCAACCCCAACAGCCGCTACTGA	480
Qy	481	CAGAGAAGCAGCAGATGTTGGACACCATCTTCAGAGATGCTCCGGAAGGAGTGCAGATC	540
Db	481	CAGAGAAGCAGCAGATGTTGGACACCATCTTCAGAGATGCTCCGGAAGGAGTGCAGATC	540
Qy	541	TAGAACAAGAAAAATGAAGGTGTGGAGAACCTCCAGAGCACTTTGATTTCAACTACAAAA	600
Db	541	TAGAACAAGAAAAATGAAGGTGTGGAGAACCTCCAGAGCACTTTGATTTCAACTACAAAA	600
Qy	601	CCCTCAAGAGCCAAAGAGACATGCGAGATCTGAATGAAACCAACAGTCTGTACACAGC	660
Db	601	CCCTCAAGAGCCAAAGAGACATGCGAGATCTGAATGAAACCAACAGTCTGTACACAGC	660
Qy	661	AGAAAGATGACAGCAGCTGGAGACCAATGCTCACACCCCTGAGACCAAGTGTGAGAAACATTG	720
Db	661	AGAAAGATGACAGCAGCTGGAGACCAATGCTCACACCCCTGAGACCAAGTGTGAGAAACATTG	720
Qy	721	TGAGTGAAGCTGGCGGGCTCTGTGTCAGCAATGAGTACGTGCGAGAAGACACTGACTGATG	780
Db	721	TGAGTGAAGCTGGCGGGCTCTGTGTCAGCAATGAGTACGTGCGAGAAGACACTGACTGATG	780
Qy	781	AAGAGCTGGCTGACTGGAAGAGGGGGCCAGAGATGCGGTGCATCTGGAGGCCCTCCCAACA	840
Db	781	AAGAGCTGGCTGACTGGAAGAGGGGGCCAGAGATGCGGTGCATCTGGAGGCCCTCCCAACA	840
Qy	841	TCTGCTGGACCGCTCTGGAAGACGTGATACCTCATTAGAGAAATTCAACTTCAGACC	900
Db	841	TCTGCTGGACCGCTCTGGAAGACGTGATACCTCATTAGAGAAATTCAACTTCAGACC	900
Qy	901	GCCAAACAATTAGAAGACTGAGAGAGCTGCAGCAGAAAGTGTCTTCAAGGGCGACCTTA	960
Db	901	GCCAAACAATTAGAAGACTGAGAGAGCTGCAGCAGAAAGTGTCTTCAAGGGCGACCTTA	960
Qy	961	TCGTGCACACCGGCCCATCATGCTGGAGGAGAGATCTGTGACGCTTTCACAAACTTAATGA	1020
Db	961	TCGTGCACACCGGCCCATCATGCTGGAGGAGAGATCTGTGACGCTTTCACAAACTTAATGA	1020
Qy	1021	AGAGTGCCTTGTGGTGGAGGGGAGGCGCTGCATGCCCCATGCAACCCGAGCCGCTTAG	1080
Db	1021	AGAGTGCCTTGTGGTGGAGGGGAGGCGCTGCATGCCCCATGCAACCCGAGCCGCTTAG	1080
Qy	1081	TCATCAAGACGTGTTCCAGTTTACACAGAAAGTCAGGTGCTGTGTCAAATTTCTGAGT	1140
Db	1081	TCATCAAGACGTGTTCCAGTTTACACAGAAAGTCAGGTGCTGTGTCAAATTTCTGAGT	1140
Qy	1141	TGAATTATACCTTAAAAATTAAAGTGTGCATGTATAAAGACTCTGGGGAGTGTGCTGCC	1200
Db	1141	TGAATTATACCTTAAAAATTAAAGTGTGCATGTATAAAGACTCTGGGGAGTGTGCTGCC	1200
Qy	1201	TCAGAGGGTCTCGGAAATTTTACATTTCTGGGGCAGAAACAAAAAGTGATGAACATGGAGG	1260
Db	1201	TCAGAGGGTCTCGGAAATTTTACATTTCTGGGGCAGAAACAAAAAGTGATGAACATGGAGG	1260
Qy	1261	AGCTTAACACGCGACCGCTGTCTCGACAGTTTCAAGCAGCTGACCCCTTAGGGAGCAGAAAT	1320
Db	1261	AGCTTAACACGCGACCGCTGTCTCGACAGTTTCAAGCAGCTGACCCCTTAGGGAGCAGAAAT	1320
Qy	1321	GTGGGAATGTGGGGCGGTGCCAAATTTGTATGCTCTTGAATCGTGAAGTGGAGGTGCACCC	1380
Db	1321	GTGGGAATGTGGGGCGGTGCCAAATTTGTATGCTCTTGAATCGTGAAGTGGAGGTGCACCC	1380
Qy	1381	TGATACACCTTGAGACTGTAGGTGTACCAACCAAGGCTCAAGATTGACCTAGAACCCACT	1440
Db	1381	TGATACACCTTGAGACTGTAGGTGTACCAACCAAGGCTCAAGATTGACCTAGAACCCACT	1440

Db 1381 TGATCACCCTTCGAGACTGAGGTGTACCAAGCCCTCAAGATTGACCTAGAGACCCTACT 1440

Qy 1441 CCTTCAGATTGTTGGTATCTCCAAACATCTGTGATGCCAAATGCTTGGGCATCAATCC 1500

Db 1441 CCTTCAGATTGTTGGTATCTCCAAACATCTGTGATGCCAAATGCTTGGGCATCAATCC 1500

Qy 1501 TGTGTATTAACATGCTCTACCAATTAACCCCAAGAGCTGTAATCTTCTACTAAGCCGCCAA 1560

Db 1501 TGTGTATTAACATGCTCTACCAATTAACCCCAAGAGCTGTAATCTTCTACTAAGCCGCCAA 1560

Qy 1561 TTGGAACCTGGAGCAAGTGGAGGCTGCTACAGCTGGAGTTCTGCTCCACCAAGC 1620

Db 1561 TTGGAACCTGGAGCAAGTGGAGGCTGCTACAGCTGGAGTTCTGCTCCACCAAGC 1620

Qy 1621 GAGGGCTGAGCATGAGCAGCTGACAAAGCTGCTGAGAAAGCTCTTAGAGGCTGTGTGA 1680

Db 1621 GAGGGCTGAGCATGAGCAGCTGACAAAGCTGCTGAGAAAGCTCTTAGAGGCTGTGTGA 1680

Qy 1681 ACTACTCAGGGTGTCAATGATGAGGCTAAATCTGCAAAAGAAACATGGCTGCCAAG 1740

Db 1681 ACTACTCAGGGTGTCAATGATGAGGCTAAATCTGCAAAAGAAACATGGCTGCCAAG 1740

Qy 1741 GCTTCTCCTTCTGGGTCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTTGG 1800

Db 1741 GCTTCTCCTTCTGGGTCTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTTGG 1800

Qy 1801 CCCCTTGGAAAGAGGTGACATCGGTTTCATCAGCAAGAGGAGGAGGCGGCATCC 1860

Db 1801 CCCCTTGGAAAGAGGTGACATCGGTTTCATCAGCAAGAGGAGGAGGCGGCATCC 1860

Qy 1861 TAAACACAAAGCCCCCGGGACCTTCTACTGCGCTTCAGCGAGAGCAGCAAAAGAGAG 1920

Db 1861 TAAACACAAAGCCCCCGGGACCTTCTACTGCGCTTCAGCGAGAGCAGCAAAAGAGAG 1920

Qy 1921 GGGTCACTTCTACTGGTGGTGAAGAGCATCACTGCGCAACCCAGATCCAGTCTGTAG 1980

Db 1921 GGGTCACTTCTACTGGTGGTGAAGAGCATCACTGCGCAACCCAGATCCAGTCTGTAG 1980

Qy 1981 AGCCATACACCAAGCAGAGCTGAAACAACATGATTTGTGTAATCATATGAGGCTATA 2040

Db 1981 AGCCATACACCAAGCAGAGCTGAAACAACATGATTTGTGTAATCATATGAGGCTATA 2040

Qy 2041 AGATCATGATGCGACCAACATCTCGTGTCTCCACTTGTACTTACCTTACCCCGCATTC 2100

Db 2041 AGATCATGATGCGACCAACATCTCGTGTCTCCACTTGTACTTACCTTACCCCGCATTC 2100

Qy 2101 CCAAGGAGAGGCAATTTGGAAGTACTGTAGGCCGAGAGCAGAGACCCCGAAGCG 2160

Db 2101 CCAAGGAGAGGCAATTTGGAAGTACTGTAGGCCGAGAGCAGAGACCCCGAAGCG 2160

Qy 2161 ACCCAGCTAGTGTGCTGCCGTACTGAAAGACCAAGTTCATCTGTGTGACACCAAGCACT 2220

Db 2161 ACCCAGCTAGTGTGCTGCCGTACTGAAAGACCAAGTTCATCTGTGTGACACCAAGCACT 2220

Qy 2221 GCAGCAATACCATTTGACCTGGCGATGTGCCCGACCTTTAGATTGATGATGAGCTTGG 2280

Db 2221 GCAGCAATACCATTTGACCTGGCGATGTGCCCGACCTTTAGATTGATGATGAGCTTGG 2280

Qy 2281 GAAATTAACGGTGAAGGTCTAGCCCTCAGCAGAGAGGAGCTTGTAGTGGCTCAAGTTTG 2340

Db 2281 GAAATTAACGGTGAAGGTCTAGCCCTCAGCAGAGAGGAGCTTGTAGTGGCTCAAGTTTG 2340

Qy 2341 ACATGATCTACCTCGAGTGTGTCTACTCCCATGTGTGAGAGCTGAAACCAAAAGCT 2400

Db 2341 ACATGATCTACCTCGAGTGTGTCTACTCCCATGTGTGAGAGCTGAAACCAAAAGCT 2400

Qy 2401 GCAGAGAGTACTTGAACACCTGCCCCGTCTCAACCCCTAAGCAGAGCCGAACCCCATTA 2460

Db 2401 GCAGAGAGTACTTGAACACCTGCCCCGTCTCAACCCCTAAGCAGAGCCGAACCCCATTA 2460

Qy 2461 TCGTCTGAACCTCTAATCTTGTGTGCTCAGATTTTTTTTTTAAATTTCTACTTCTGCT 2520

Db 2461 TCGTCTGAACCTCTAATCTTGTGTGCTCAGATTTTTTTTTTAAATTTCTACTTCTGCT 2520

Qy 2521 ATCTTTGGCAGATCTGGGCACTTTTAAAGAGAGAAATGAGTACGTGTGGTATTAAC 2580

Db 2521 ATCTTTGGCAGATCTGGGCACTTTTAAAGAGAGAAATGAGTACGTGTGGTATTAAC 2580

Qy 2581 TGTATGTAAAGAGAGAGAGCTCTGAGTCTGGGATGTGGGCTGAGACAGAAAGAGAGC 2640

Db 2581 TGTATGTAAAGAGAGAGAGCTCTGAGTCTGGGATGTGGGCTGAGACAGAAAGAGAGC 2640

Qy 2641 AAAAGGAGAACCTCTCTGCTGCGCCGCTGCGCTCTCTTTTTCAGACGCTCGGGGCTTGG 2700

Db 2641 AAAAGGAGAACCTCTCTGCTGCGCCGCTGCGCTCTCTTTTTCAGACGCTCGGGGCTTGG 2700

Qy 2701 TTTTAAACAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760

Db 2701 TTTTAAACAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760

Qy 2761 ACCCATTTCTGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820

Db 2761 ACCCATTTCTGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820

Qy 2821 TAGACTTAAGCCAGAGAGTCTCTTTAAATTAAGAGAGAGAGAGAGAGAGAGAGAG 2869

Db 2821 TAGACTTAAGCCAGAGAGTCTCTTTAAATTAAGAGAGAGAGAGAGAGAGAGAGAG 2869

RESULT 6

LOCUS BC003806 2964 bp mRNA ROD 12-JUL-2001

DEFINITION Mus musculus, Signal transducer and activator of transcription 3, clone MGC:6089 IMAGE:3593588, mRNA, complete cds.

ACCESSION BC003806

VERSION BC003806.1 GI:13277851

KEYWORDS MGC.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE Eukarya: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2964)

AUTHORS Strausberg, R.

JOURNAL Direct Submission

Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT Contact: MGC help desk

Email: gcabs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcgdpaxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES

source

1..2964

Location/Qualifiers

/organism="Mus musculus"

/db_xref="LocusID:25125"

/db_xref="taxon:10090"

/clone="MGC:6089 IMAGE:3593588"

/tissue_type="Mammary tumor, Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."

/clone_id="NCI_CGAP_Mam1"

/lab_host="DH10B"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 476715.

CDS

/note="Vector: PCMV-SPORT6"
168..2480
/codon_start=1
/product="Signal transducer and activator of transcription
3"

/protein_id="AAH03806.1"
/db_xref="GI:13277852"

/translation="MAQMNQLOLDTRYLEQHLQYSDSPFMEIRFLAPWIESQDMA
YAASKESHATLVEHNLGLGEIDQOYRFLQSNVLYQNLNRIRIKOFLDSRYLEKMEIA
RIVARCLMEESRLDTATAOOGGQANHTPAAVTEKQOOLEQHLQDVRKRVODLEQ
KMKVVENIODDFENYKTKLSQGMODLNQSNQVTRQKQKQLEQMLTALDOMRSIV
SELAGILAMEYVOKLTIDELADMKRROOACIGSPNIDLEMLNITSLAESQD
TROOIKLEPELOQVSKYKQPIVQHRPMLERLVELEFRLNLSAFVVERQPCQMHEHD
RPLVKTGVQFTTKVLLVLPPELNTOLKIKVCIDKRSQVYALRGRKRNILSTNPK
VMNEESNNSLSAEFKHLLREQRNGNGRANCDASLIVTEHLHLTFTEVTHQGL
KIDLETSLPVLVVISNICOMPNAWASILWMLNMLNPNKNVNEFKPIIGTMDQVAYEL
SWQPSSTTKRLGLSLEQLTLAEKLLGPGVANSQCOITMAKFCXENMAGKGSFPMWLD
NIIDLVKYLIALNEGYIMGFISKERERAILSTPKPGFTFLRESSEKGEQVFTVY
EKDISGKTQIOVSVEPYTKOOLNNMSFAEIMGYKIMATNTILVPLVLYLDPIKEEA
FGKYCRPESQEHPEADPGSAPRYLTKTKPFICYTPTTCSNTIDLPSPPTLDSLMQFGNN
GEGEPASAGGFESLTFMDLTSFCATSPM"

BASE COUNT 764 a 782 c 807 g 611 t
ORIGIN

Query Match 96.5%: Score 2767.8; DB 10; Length: 2964;

Best Local Similarity 99.2%: Pred. No. 0; Mismatches 17; Indels 7; Gaps 5;

Matches 2835; Conservative 0; Mismatches 17; Indels 7; Gaps 5;

QY 13 CAGCCGCGCAGTCCGGC--TCAGCCCGAGACAGTCGAGACCCG13ACTGCGAGAGAT 70
DB 110 CAGCCGCGCAGTCCGGCAGCTCGAGCCGAGACAGTCGAGACCCGCTGACTGCGAGAGAT 169
QY 71 GGCTAGTGAACCGCTGCGACGACTGACACGCTACCTGAGCGAGCTGCGACAGCT 130
DB 170 GGCTAGTGAACCGCTGCGACGACTGACACGCTACCTGAGCGAGCTGCGACAGCT 229
QY 131 GTACAGCAGACGTTCCCATATGAGCTGGGCGAGTTCTCGGCACTTGGATTGAGAGTCA 190
DB 230 GTACAGCAGACGTTCCCATATGAGCTGGGCGAGTTCTCGGCACTTGGATTGAGAGTCA 289
QY 191 AACAGTGGCATATGACGACGAGCAAGAGTCACATGCCAGCTGGTGTTCATATCTCT 250
DB 290 AACAGTGGCATATGACGACGAGCAAGAGTCACATGCCAGCTGGTGTTCATATCTCT 349
QY 251 GGGTAAATTTGACGACATATAGCCGATTTCTGCAAGAGTCGAATGTCTTATCAGCA 310
DB 350 GGGTAAATTTGACGACATATAGCCGATTTCTGCAAGAGTCGAATGTCTTATCAGCA 409
QY 311 CAACCTTGAAGATCAGCAATGTTCTGACAGCAGTATCTTGAAGAGCCAAATGGAAT 370
DB 410 CAACCTTGAAGATCAGCAATGTTCTGACAGCAGTATCTTGAAGAGCCAAATGGAAT 469
QY 371 TGCCTGATCGTGGCCGATGCTGTTGGGAAGAGTCTGCTCTCCACAGCGGACCCAC 430
DB 470 TGCCTGATCGTGGCCGATGCTGTTGGGAAGAGTCTGCTCTCCACAGCGGACCCAC 529
QY 431 GGCAGCCGACGAAGGGGGCCAGGCCAACCCCAACAGCCGCGCTGATGACAGAGA 490
DB 530 GGCAGCCGACGAAGGGGGCCAGGCCAACCCCAACAGCCGCGCTGATGACAGAGA 589
QY 491 GGAGATGTTGGAGCAGCATCTTCAGGATGTCGGAAAGCAGTGCAGGATCTAGAAAGAA 550
DB 590 GGAGATGTTGGAGCAGCATCTTCAGGATGTCGGAAAGCAGTGCAGGATCTAGAAAGAA 649
QY 551 AATGAAGTGTGAGAAACCTCCAGAGCAGTTTGAATTTCACTACAAAACCTCAAGAG 610
DB 650 AATGAAGTGTGAGAAACCTCCAGAGCAGTTTGAATTTCACTACAAAACCTCAAGAG 709
QY 611 CCAAGGAGCATGCGAGATCTGAATGGAACAACAGCTGTGACGACAGAGAAGTGA 670
DB 710 CCAAGGAGCATGCGAGATCTGAATGGAACAACAGCTGTGACGACAGAGAAGTGA 769

QY 671 GCAGCTGAACAGATGCTCAAGCCCTGGACCAAGATGGGGAAGCATTTGTAGTACGT 730
DB 770 GCAGCTGAACAGATGCTCAAGCCCTGGACCAAGATGGGGAAGCATTTGTAGTACGT 829
QY 731 GCGCGGCTCTTGTACGCAATGAGTACGTGACAGAGACATGATGAGAGAGTGC 790
DB 830 GCGCGGCTCTTGTACGCAATGAGTACGTGACAGAGACATGATGAGAGAGTGC 889
QY 791 TGACTGGAAGAGCGGCCGAGAGATCGCGTGCATGGAGAGCCCTCCCAACATCTCCCGGA 850
DB 890 TGACTGGAAGAGCGGCCGAGAGATCGCGTGCATGGAGAGCCCTCCCAACATCTCCCGGA 949
QY 851 CCGTGTGAAAACCTGATTAATCTTCAATTCAGCAATCTCAAGGCCACCTATCTGACACA 910
DB 950 CCGTGTGAAAACCTGATTAATCTTCAATTCAGCAATCTCAAGGCCACCTATCTGACACA 1009
QY 911 TAAGAACTGGAGAGAGCTGACAGCAGAAAGTCTCTCAAGGGCAGCCCTATCTGACACA 970
DB 1010 TAAGAACTGGAGAGAGCTGACAGCAGAAAGTCTCTCAAGGGCAGCCCTATCTGACACA 1069
QY 971 CCGGCCATCTGTCGAGGAGAGATCGTGAGCTGTTCAGAACTTAATGAAGAGTGCCTT 1030
DB 1070 CCGGCCATCTGTCGAGGAGAGATCGTGAGCTGTTCAGAACTTAATGAAGAGTGCCTT 1129
QY 1031 CCGTGTGAGAGCGGACAGCCCTGCATGACCATGACACCCGAGCCGCTTATGATCAAGAC 1090
DB 1130 CCGTGTGAGAGCGGACAGCCCTGCATGACCATGACACCCGAGCCGCTTATGATCAAGAC 1189
QY 1091 TCGTGTCCAGTTTACACAGAAAGTACGTTGCTGTCAAATTTCTGTGATTGAATATCA 1150
DB 1190 TCGTGTCCAGTTTACACAGAAAGTACGTTGCTGTCAAATTTCTGTGATTGAATATCA 1249
QY 1151 GCTTAAATTTAAGTGTGATTAAGACACTCTGGGAGTCTTGTGCTGCTCAGAGGCTC 1210
DB 1250 GCTTAAATTTAAGTGTGATTAAGACACTCTGGGAGTCTTGTGCTGCTCAGAGGCTC 1309
QY 1211 TCGGAATTTAATCAATTTGCGGACGAGACAAAGTATGAACATGAGAGCTTAACAA 1270
DB 1310 TCGGAATTTAATCAATTTGCGGACGAGACAAAGTATGAACATGAGAGCTTAACAA 1369
QY 1271 GGGCACTGTGTCGAGAGTTCAAGCAGCTGACCTTGGAGAGAGAGATGTGGAAATGG 1330
DB 1370 GGGCACTGTGTCGAGAGTTCAAGCAGCTGACCTTGGAGAGAGAGATGTGGAAATGG 1429
QY 1331 AGCCGTCGCAATTTGATGCTCTCTTATGATGTCGACTGAGAGAGCTGACATCACCTT 1390
DB 1430 AGCCGTCGCAATTTGATGCTCTCTTATGATGTCGACTGAGAGAGCTGACATCACCTT 1489
QY 1391 CGAGACTGAGTGTACACCAAGAGGCTCAAGATTGACCTAGAGACCACTCTTGCCAGT 1450
DB 1490 CGAGACTGAGTGTACACCAAGAGGCTCAAGATTGACCTAGAGACCACTCTTGCCAGT 1540
QY 1451 TGTGTGATCTCCAAATCTGTACAGATGCAAAATCTTGGGATCAATCTGTGTATTA 1510
DB 1550 TGTGTGATCTCCAAATCTGTACAGATGCAAAATCTTGGGATCAATCTGTGTATTA 1609
QY 1511 CATGCTGACCAATTAACCCCAAGAGTGAACCTTTCACTAAAGCCGCCAATTTGAACCTG 1570
DB 1610 CATGCTGACCAATTAACCCCAAGAGTGAACCTTTCACTAAAGCCGCCAATTTGAACCTG 1669
QY 1571 GGACCAAGTGGCGGAGGTCGCTGAGTGGAGTTCGTCACACCAAGAGCAAGGAGCTGAG 1630
DB 1670 GGACCAAGTGGCGGAGGTCGCTGAGTGGAGTTCGTCACACCAAGAGCAAGGAGCTGAG 1729
QY 1631 CATCGAGCAGCTGACACAGCTGGCTGAGAGAGCTCTAGGGCTTGTGTGAATCTACAGG 1690
DB 1730 CATCGAGCAGCTGACACAGCTGGCTGAGAGAGCTCTAGGGCTTGTGTGAATCTACAGG 1789
QY 1691 GTGTGAGATCACATGGGCTAAATTTCTGCAAGAAAGAACTGGCTGCAAGAGGCTTCTCTT 1750
DB 1790 GTGTGAGATCACATGGGCTAAATTTCTGCAAGAAAGAACTGGCTGCAAGAGGCTTCTCTT 1849
QY 1751 CTGGGCTGTGCTAGACAAATATCATGACCTTGTGAAAAAATATATCTTGGCCCTTTGAAA 1810

Db	1850	CTGGGTCTGGCTAGCAAAATATCTGCACCTTGTGAAAAAGTATCTTGGCCCTTTGGAA	1909
Qy	1811	TGAAGGGTACATATGGGTTTCATGACGAAGAGCGGGAGCGGCATCTTAAGCACAA	1870
Db	1910	TGAAGGGTACATATGGGTTTCATGACGAAGAGCGGGAGCGGCATCTTAAGCACAA	1969
Qy	1871	GCCCCGGGACCTTCTCTACTGGCGTTTCAGACGACGAAAGAAAGAGGGGTCACTT	1930
Db	1970	GCCCCGGGACCTTCTCTACTGGCGTTTCAGACGACGAAAGAAAGAGGGGTCACTT	2029
Qy	1931	CACTTGGGTGAAAGAGACATCATGTGGCAAGACCCAGATCCAGTCTGTAGAGCCATAC	1990
Db	2030	CACTTGGGTGAAAGAGACATCATGTGGCAAGACCCAGATCCAGTCTGTAGAGCCATAC	2089
Qy	1991	CAAGCAGACTGAAACAACATGTCTATTGTCTGAAATCATCATGGGCTATTAAGATCATGA	2050
Db	2090	CAAGCAGACTGAAACAACATGTCTATTGTCTGAAATCATCATGGGCTATTAAGATCATGA	2149
Qy	2051	TGCAGCCAAATCCTGGTGTCTCCACTTGTCTACCTCTACCCCGCAATYCCCAAGAGGA	2110
Db	2150	TGCAGCCAAATCCTGGTGTCTCCACTTGTCTACCTCTACCCCGCAATYCCCAAGAGGA	2209
Qy	2111	GGCATTTGGAAATAGTACTGTATAGGCCCGAAGACCAGAGACACCCCGAAGCCAGCTAG	2170
Db	2210	GGCATTTGGAAATAGTACTGTATAGGCCCGAAGACCAGAGACACCCCGAAGCCAGCTAG	2269
Qy	2171	TGCTGCCCCGACTTGAAGACCAAGTTCACTGTGTGACACCAAGACCTGAGCAATAC	2230
Db	2270	TGCTGCCCCGACTTGAAGACCAAGTTCACTGTGTGACACCAAGACCTGAGCAATAC	2329
Qy	2231	CATTGACCTGCCAATGTGCCCCCGCACTTTAGATTCAATTGATGCAGTTTGAATAACGG	2290
Db	2330	CATTGACCTGCCAATGTGCCCCCGCACTTTAGATTCAATTGATGCAGTTTGAATAACGG	2389
Qy	2291	TGAAGGTGTGAGGCGCTCAGACAGAGAGGAGTTGATGTCGTACAGTTTGACATGATCACT	2350
Db	2390	TGAAGGTGTGAGGCGCTCAGACAGAGAGGAGTTGATGTCGTACAGTTTGACATGATCACT	2449
Qy	2351	GACCTGGAGTGTGTACCTCTCCCCCATGTGTAGAGCTGAAACAGAAAGCTGCAGAGAGCT	2410
Db	2450	GACCTGGAGTGTGTACCTCTCCCCCATGTGTAGAGCTGAAACAGAAAGCTGCAGAGAGCT	2509
Qy	2411	GACTTGAGACACCTGCCCCGTGCTCCACCCCTTAAGACGCCGAACCCATATGTCGTGAA	2470
Db	2510	GACTTGAGACACCTGCCCCGTGCTCCACCCCTTAAGACGCCGAACCCATATGTCGTGAA	2569
Qy	2471	CTCTCACTTGTGTGTCCAGATTTTTTTTTTTAAATTCCTACTCTGCTATCTTTGGC	2530
Db	2570	CTCTCACTTGTGTGTCCAGATTTTTTTTTTTAAATTCCTACTCTGCTATCTTTGGC	2629
Qy	2531	AATCTGGGCACTTTTAAAGAGAGAAATGTGAGTGTGGGTGATAACTGTATGTGA	2590
Db	2630	AATCTGGGCACTTTTAAATAGAGAAATGTGAGTGTGGGTGATAACTGTATGTGA	2689
Qy	2591	AGAGAGAG--ACCTGTAGCTGTGGGATGGGCTGTGAGACAGAAAGGAGCAAGAGGGA	2649
Db	2690	AGAGAGAGCACTCTAGTCTGGGATGGGCTGTGAGAGCAGAAAGGAG--CAAGGGGA	2747
Qy	2650	CACCTCTGTCTTCCGCCCTGTCCCTCTTTTTCAGACAGCTGGGGGTTGTGTTAGAC	2709
Db	2748	CACCTCTGTCTTCCGCCCTGTCCCTCTTTTTCAGACAGCTC--GGGGTTGTTGTTAGAC	2808
Qy	2710	AAGTGTCTCTGGGTGCCATGTGGCTACTGTGTGGCCCACTGTGTGAGCTATACCCATTG	2768
Db	2807	AAGTGTCTCTGGGTGCCCAATGGGATCTGTGTGCCCACTGTGTGAGCTATACCCAGAGC	2866
Qy	2770	TGGGAACCTCGGTGTGCACCTTTCACACTTGTATATTCACATAGAACTAGAGACTTA	2829
Db	2867	TGGGAACCTCGGTGTGCACCTTTCACACTTGTATATTCACATAGAACTAGAGACTTA	2928
Qy	2830	GCCAGAGAGTTCTCTTTAAATTTAAAAA	2868

DB	2927	GCCCA-GAGGTCCTCTTAATTAATAAAAAAAAAAAAAA	2964
RESULT	7		
LOCUS	MMU30709	2924 bp	17-JUN-1996
DEFINITION	Mus musculus Stat3b mRNA, complete cds.		
ACCESSION	U30709		
VERSION	U30709.1	GI:1166505	
KEYWORDS			
SOURCE			
ORGANISM	house mouse.		
REFERENCE	Mus musculus.		
AUTHORS	Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 2924)		
JOURNAL	Schaefer,T.S., Sanders,L.K. and Nathans,D.		
MEDLINE	Cooperative transcriptional activity of Jun and Stat3 beta, a short form of Stat3		
REFERENCE	Proc. Natl. Acad. Sci. U.S.A. 92 (20), 9097-9101 (1995)		
AUTHORS	2 (bases 1 to 2924)		
TITLE	Schaefer,T.S., Sanders,L.K. and Nathans,D.		
JOURNAL	Submitted (30-JUN-1995) Timothy S. Schaefer, Mol. Biol. & Genetics, Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205, USA		
REFERENCE	3 (bases 1 to 2924)		
AUTHORS	Caldenhoven,E., van Dijk,T.B., Solari,R., Armstrong,J., Raaijmakers,J.A.M., Lammers,J.W.J., Koenderman,L. and de Groot,R.P.		
TITLE	Stat3beta, a splice variant of transcription factor STAT3, is a dominant negative regulator of transcription		
JOURNAL	J. Biol. Chem. 271 (22), 13221-13227 (1996)		
FEATURES	9628730		
SOURCE	Location/Qualifiers		
	1..2924		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	173..2341		
	/codon_start=1		
	/product="Stat3b"		
	/protein_id="AAC52612.1"		
	/db_xref="GI:1166506"		
	/translation="MAOMNLOLODTRYLEOLHOLYSDPPEMLROFLAPIESODMAN YAASKESHATVYFHNHLAGLEIDOOYSRLQDSENVLYKHNLHROFLDLSRIEKEKMEIATA RIVACLWIEESRLQTAATAAODGGOAHPTAAVTEQOLBOHLODRVRVODLE KMKVLENLDQDFDENYKTKLKSQGMODLNGNSOYTRKMOOLEDMITLDMQMSRIV SELKVLAMEYVOFKTLDEBLADMRKQOJACIGGPNICDLRLNMTSLAESODL RTRQIKKLEELQOKKSYKGD IVQHRPELVYELFENLKSAFVVEROPCMHPDD PVLQIKGVQFTTKYRLVLRPELVYOLKIVCIDKDSGDVAALFGSKRNFILGTNTR VMNNEESNGLSAEFKHLTLRBOCRGSGANDASLIYVEELHLITFEFEYVHOGI KIDETHSLPVTYVINSNICOMRNANASILMYMNLNTRNVPFLPKPIGTMDVAAEVL SWQSSSTTKRGLSIBOLTTLAEKLGPECVNTSGCQITNAFKCKSMGKGSFVWMLD NIDLVKKYLIILAMNEGYIMGFIKSERERAILSTRKPTGTLFLRESSEKSGEGVFTVW EKDLSGKTOIOISVEPYTKOOLNNMSFAEIIINGYKIMDATNLIVSPLVLYLPDIPKEEAE FGKYCRPESOEHPENDPCSAAPYILKTRKICVTPEFDAVMK"		
BASE COUNT	756 a 764 c 801 g 603 t		
ORIGIN			
Query Match	94.5%; Score 2711.2; DB 10; Length 2924;		
Best Local Similarity	97.9%; Pred. No. 0;		
Matches 2799; Conservative	0; Mismatches 8; Indels 53; Gaps 3;		
DB	13	CAGGCCGGCACTGGGGC--TCAGCCCGGAGACACAGTCCGAGACCCCTGACTGACAGAGAT	70
DB	115	CAGGCCGGCACTGGGGCCTTCAGCCCGGAGACAGTCCGAGACCCCTGACTGACAGAGAT	174
DB	71	GAGCTAGTGAACAGCTCAGACAGTGGAGACAGCGTCGTAAGACAGCTCAGACAGT	130
DB	175	GAGCTAGTGAACAGCTCAGACAGTGGAGACAGCGTCGTAAGACAGCTCAGACAGT	234
DB	131	GTACAGCAGACAGTTCCTCCCATGAGAGCTGGCGCAGTTCCTGGCACCCTTGAGATTGAGAGTCA	190


```
Db 235 GTACAGCAGACGTTCCCATGAGCTGCGCAGTTCCTGCGACCTTGATGAGAGTCA 294
QY 191 AGACGGGCATATGAGCGCCCAAGAGTCCACATGCCAGTTGGTGTTCATTAATCTTT 250
Db 295 AGACTGGCATATGAGCGCCCAAGAGTCCACATGCCAGTTGGTGTTCATTAATCTTT 354
QY 251 GGGTGAATTTGACGACGATATAGCCGATTCCTGAGAGTCCAAATGTCCTATCACA 310
Db 355 GGGTGAATTTGACGACGATATAGCCGATTCCTGAGAGTCCAAATGTCCTATCACA 414
QY 311 CAACCTTGGAGAAATCAAGCAAGTTCTGAGAGAGTATCTTGAGAGCCAAATGGAAT 370
Db 415 CAACCTTGGAGAAATCAAGCAAGTTCTGAGAGAGTATCTTGAGAGAGCCAAATGGAAT 474
QY 371 TGCCTGATCTGTCGCGCATGCTGTGGAGAGTCTGCTCTCTGACAGCGCAC 430
Db 475 TGCCTGATCTGTCGCGCATGCTGTGGAGAGTCTGCTCTCTGACAGCGCAC 534
QY 431 GGCAGCCCAAGAGGGGCGCCAGCCACCCAGCCCGCCGTAAGTACAGAGAGCA 490
Db 535 GGCAGCCCAAGAGGGGCGCCAGCCACCCAGCCCGCCGTAAGTACAGAGAGCA 594
QY 491 GCAGATGTTGAGAGCAGCATCTCAGAGTGTCCGAAAGCAGTGTAGAGATCTAGAACGA 550
Db 595 GCAGATGTTGAGAGCAGCATCTCAGAGTGTCCGAAAGCAGTGTAGAGATCTAGAACGA 654
QY 551 AATGAAGTGTGAGAGACCTCCAGAGACATTTGATTCACTAACAAACCTCAAGAG 610
Db 655 AATGAAGTGTGAGAGACCTCCAGAGACATTTGATTCACTAACAAACCTCAAGAG 714
QY 611 CCAAGAGACATGACAGATCTGAATGGAACAAACAGTCTGTACAGAGAGAGATGCA 670
Db 715 CCAAGAGACATGACAGATCTGAATGGAACAAACAGTCTGTACAGAGAGAGATGCA 774
QY 671 GCAGCTGGAACAGATGCTCAGACCCCTGACAGACAGATGCGGAGAGAGATGTAGTAGCT 730
Db 775 GCAGCTGGAACAGATGCTCAGACCCCTGACAGACAGATGCGGAGAGAGATGTAGTAGCT 834
QY 731 GCGGGGCTCTTGTCTCAGATGAGTACCTGACAGAGACACTGATGAGAGAGCTGCG 790
Db 835 GCGGGGCTCTTGTCTCAGATGAGTACCTGACAGAGACACTGATGAGAGAGCTGCG 894
QY 791 TGACTGGAAGAGCGCGCAGAGATGCGGTGATCGGAGGCGCTCCAAATCTGCTGGA 954
Db 895 TGACTGGAAGAGCGCGCAGAGATGCGGTGATCGGAGGCGCTCCAAATCTGCTGGA 954
QY 851 CCGTGTGAAACCTGATATCACTTAGAGAGATCTCAACTTCAGACCGCCCAACAAAT 910
Db 955 CCGTGTGAAACCTGATATCACTTAGAGAGATCTCAACTTCAGACCGCCCAACAAAT 1014
QY 911 TAAGAAACTGAGAGAGCTGACGAGAAAGTGTCTACAAAGGCGACCTATCTGTGACGA 970
Db 1015 TAAGAAACTGAGAGAGCTGACGAGAAAGTGTCTACAAAGGCGACCTATCTGTGACGA 1074
QY 971 CCGGCCATGCTGAGAGAGAGTGTGAGAGTTCAGAACTTATGAGAGTGTGCTT 1030
Db 1075 CCGGCCATGCTGAGAGAGAGTGTGAGAGTTCAGAACTTATGAGAGTGTGCTT 1134
QY 1031 CCGTGTGAGAGCGCGCAGCGCTGATGCCATGCAACCGGACCGCCCTTACTGATCAAGAC 1090
Db 1135 CCGTGTGAGAGCGCGCAGCGCTGATGCCATGCAACCGGACCGCCCTTACTGATCAAGAC 1194
QY 1091 TGTGTGTCACATTTACCGAGAAAGTCAAGTGTGAGTCAAAATTTCTGAGATTGAATTATCA 1150
Db 1195 TGTGTGTCACATTTACCGAGAAAGTCAAGTGTGAGTCAAAATTTCTGAGATTGAATTATCA 1254
QY 1151 GCTTAAATTTAAAGTGTGATTAAGACTGTGGGATGTTGCTGCTCAGAGGCTC 1210
Db 1255 GCTTAAATTTAAAGTGTGATTAAGACTGTGGGATGTTGCTGCTCAGAGGCTC 1314
QY 1211 TCGGAAATTTAACTTGTGGGACAGAACACAAAGTATACATGAGAGAGTCTAACAA 1270
Db 1315 TCGGAAATTTAACTTGTGGGACAGAACACAAAGTATGAGATGAGAGAGTCTAACAA 1374

QY 1271 CCGCAGCCTGTCTGACAGAGTTCAGACACCTGAGCCCTTAGGGAGCAGAGATGTGGAAATGG 1330
Db 1375 CCGCAGCCTGTCTGACAGAGTTCAGACACCTGAGCCCTTAGGGAGCAGAGATGTGGAAATGG 1434
QY 1331 AGGCCGTGCAATTTGATGTGCTCCTCTGATGATGATGAGAGAGTGTGACCTGATCACCCTT 1390
Db 1435 AGGCCGTGCAATTTGATGTGCTCCTCTGATGATGATGAGAGAGTGTGACCTGATCACCCTT 1494
QY 1391 CGAGACTAGGTTGATACCAACCAAGGCTTCAGATTGACTTAAGAGCCCACTCTGCGCAGT 1450
Db 1495 CGAGACTAGGTTGATACCAACCAAGGCTTCAGATTGACTTAAGAGCCCACTCTGCGCAGT 1554
QY 1451 TGTGTGATTTCCCAACATCTGTGACATGCAAAATCTTGGGCATCAATCTGTGATATA 1510
Db 1555 TGTGTGATTTCCCAACATCTGTGACATGCAAAATCTTGGGCATCAATCTGTGATATA 1614
QY 1511 CATGCTGACCAATTAACCCCAAGAGAGTGAATCTTCACTAAAGCCCAATTGGAACCTG 1570
Db 1615 CATGCTGACCAATTAACCCCAAGAGAGTGAATCTTCACTAAAGCCCAATTGGAACCTG 1674
QY 1571 GGACCAAGTGGCGGAGGTGCTCAGCTGCGAGTTCGTGTCACACCAAGCGAGGCTGAG 1630
Db 1675 GGACCAAGTGGCGGAGGTGCTCAGCTGCGAGTTCGTGTCACACCAAGCGAGGCTGAG 1734
QY 1631 CATGAGACAGTGAACACGCTGCTGAGAAAGCTCTAGGCGCTGTGTGAATCTACAGG 1690
Db 1735 CATGAGACAGTGAACACGCTGCTGAGAAAGCTCTAGGCGCTGTGTGAATCTACAGG 1794
QY 1691 GTGTGATGATCATGATGGGCTTAATTTCTGCAAAAGAAACATGGCTGGCAAGGGCTTCTCT 1750
Db 1795 GTGTGATGATCATGATGGGCTTAATTTCTGCAAAAGAAACATGGCTGGCAAGGGCTTCTCT 1854
QY 1751 CTGGTCTGTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTTGGCCCTTGGAA 1810
Db 1855 CTGGTCTGTGCTAGACAAATATCATGACCTTGTGAAAAAGTATATCTTGGCCCTTGGAA 1914
QY 1811 TGAAGGTTACATCATGTGGTTTCATGCAAGAGCGGAGCGGCGCATCTTAAGACAAA 1870
Db 1915 TGAAGGTTACATCATGTGGTTTCATGCAAGAGCGGAGCGGCGCATCTTAAGACAAA 1974
QY 1871 GCGCCCGGCGACCTTCTCTAGCGCTTACGCGAGCAGACCAAGAGAGGGGTCACTTT 1930
Db 1975 GCGCCCGGCGACCTTCTCTAGCGCTTACGCGAGCAGACCAAGAGAGGGGTCACTTT 2034
QY 1931 CACTTGGTGTGAAAAAGACATCATGTGGCAAGACCCAGATCTGTGTAGACCATACAC 1990
Db 2035 CACTTGGTGTGAAAAAGACATCATGTGGCAAGACCCAGATCTGTGTAGACCATACAC 2094
QY 1991 CAAGCAGAGCTGAACACATGCTGTAATGTAATCATGAGGCTATTAAGTATGGA 2050
Db 2095 CAAGCAGAGCTGAACACATGCTGTAATGTAATCATGAGGCTATTAAGTATGGA 2154
QY 2051 TGCAGCAACATCTGCTGTCTCAGCTGTACCTTACCCCGCATTTCCAAAGAGGA 2110
Db 2155 TGCAGCAACATCTGCTGTCTCAGCTGTACCTTACCCCGCATTTCCAAAGAGGA 2214
QY 2111 GGCATTTGGAAGTACTGTAGAGCGCGCAGAGCGGAGCGCCGAGCGGACGAGGTAG 2170
Db 2215 GGCATTTGGAAGTACTGTAGAGCGCGCAGAGCGGAGCGCCGAGCGGACGAGGTAG 2274
QY 2171 TGTGTGCGCGTACCTGAAGACCAAGTTCATCTGTGTAGACCAAGCAAGCTGTGACGAATAC 2230
Db 2275 TGTGTGCGCGTACCTGAAGACCAAGTTCATCTGTGTAGACCAAGCTGTGACGAATAC 2316
QY 2231 CATTTGACCTGCGCATGTCCCGCGCACCTTATGATTTGATGATGATTTGAAATTAAGG 2290
Db 2317 CATTTGACCTGCGCATGTCCCGCGCACCTTATGATTTGATGATGATTTGAAATTAAGG 2344
QY 2291 TGAAGGTTGAGAGCGCTCAGCAGAGAGGAGTGTGAGTGCCTCAGCTTGGATGATGATCT 2350
Db 2345 TGAAGGTTGAGAGCGCTCAGCAGAGAGGAGTGTGAGTGCCTCAGCTTGGATGATGATCT 2404
```

OY	2351	GACCTCGAGTGTGCTACTCTCCCATCTGAGGAGCTAAACCAGAAGCTGCAGACGT	2410
OY	2405	GACCTCGAGTGTGCTACTCTCCCATCTGAGGAGCTAAACCAGAAGCTGCAGACGT	2464
OY	2411	GACTTGACAACATCTGCCCGCTGCTCACCCCTTAACAGACGGACCCCATATCTGTGAAA	2470
OY	2465	GACTTGACAACATCTGCCCGCTGCTCACCCCTTAACAGACGGACCCCATATCTGTGAAA	2524
OY	2471	CTCCTCACTTTGTGGTTCACATTTTTTTTTTAAATTCTGACTCTGCTATCTTTGGGC	2530
Db	2525	CTCCTCACTTTGTGGTTCACATTTTTTTTTTAAATTCTGACTCTGCTATCTTTGGGC	2584
OY	2531	AATCTGGGCACTTTTTTAAAGAGAGAATAAGTAGTGTGGGTGAATACTGTATGTAA	2590
Db	2585	AATCTGGGCACTTTTTTAAAGAGAGAATAAGTAGTGTGGGTGAATACTGTATGTAA	2644
OY	2591	AGAGAGAGAG-ACCTGTAGCTGTGGGGATGGGGCTAAGAGCAGAAAGGAGCCAAAGGGGA	2649
Db	2645	AGAGAGAGAGCACTTGTAGCTGTGGGGATGGGGCTAAGAGCAGAAAGGAGCCAAAGGGGA	2704
OY	2650	CACCTCTGTCTCTGCCCCCCTGCCCCTCTCTTTTTCAGCAGCTGGGGGTTGGTTTGAAC	2709
Db	2705	CACCTCTGTCTCTGCCCCCCTGCCCCTCTCTTTTTCAGCAGCTGGGGGTTGGTTTGAAC	2764
OY	2710	AAGTGCCTCTGCTGGTCCCATGGCTAACCTGCTGCTGCTGAGCTGATACCCCATTC	2769
Db	2765	AAGTGCCTCTGCTGGTCCCATGGCTAACCTGCTGCTGAGCTGATACCCCATTC	2824
OY	2770	TGGGAAGTCCCTGGCTGTGACTTTAACCTTGCTAATATTCACATAGAAAGCTAGGACTPA	2829
Db	2825	TGGGAAGTCCCTGGCTGTGACTTTAACCTTGCTAATATTCACATAGAAAGCTAGGACTPA	2884
OY	2830	GCCCAAGAGGTTCTCTTTAAATTAAAAAATAAAAAAAAA	2869
Db	2885	GCCCAAGAGGTTCTCTTTAAATTAAAAAATAAAAAAAAA	2924
RESULT	8		
LOCUS	MMU08378	2924 bp	07-DEC-1994
DEFINITION	Mus musculus acute phase response factor (APRF) mRNA, complete cds.		
VERSION	U08378		
KEYWORDS	U08378.1 GI:473889		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Raz,R., Durbin,J.E. and Levy,D.E.		
TITLE	Acute phase response factor and additional members of the interferon-stimulated gene factor 3 family integrate diverse signals from cytokines, interferons, and growth factors		
JOURNAL MEDLINE REFERENCE	J. Biol. Chem. 269, 24391-24395 (1994)		
AUTHORS TITLE	Levy,D.E.		
JOURNAL	Direct Submission		
FEATURES	Submitted (05-APR-1994) David E. Levy, Pathology, New York University School of Medicine, 550 First Avenue, New York, NY 10016, USA		
SOURCE	Location/Qualifiers		
	1..2924	/organism="Mus musculus"	
	/db_xref="taxon:10090"		
	/clone="PBS30"		
	/clone_lib="Unizap library (Stratagene)"		
	/tissue_type="brain"		
	/dev_stage="newborn"		
	1..146		
	147..2456		
	/note="acute phase response factor"		
	/codon_start=1		

[illegible]

Qy	728	CTGGCGGGCCCTCTTGTGACGCAATGAGTAGTACGTGCGAGAAAGACACTGACTGATGAAAGACT	787
Db	806	GCTGGCGGGGCCCTTGTGAGCAATGAGTAGTGTGCGAGAAAGACACTGACTGATGAAAGACT	865
Qy	788	GGCTGACTGGAAGAGGGCGCGAGAGATGCGGTGCATGCGGAGGGCCCTCCCAACTGTGCGCT	847
Db	866	GGCTGACTGGAAGAGGGCGCGACAGATGCGGTGCATGCGGAGGGCCCTCCCAACTGTGCGCT	925
Qy	848	GGACCGCTTGGAAAACTGATTAACCTTCATTAGCAGAAATCTCAACTTCACACCCGGCCCAAC	907
Db	926	GGACCGCTCTGGAAAACTGATTAACCTTCATTAGCAGAAATCTCAACTTCAGACCCGGCCCAAC	985
Qy	908	AATTAAAGAAACTGAGAGAGACTCGACAGAGAAAGTGTCTCAAGAGGGGACCCCTATGCTGCA	967
Db	986	AATTAAAGAAACTGAGAGAGACTCGACAGAGAAAGTGTCTCAAGAGGGGACCCCTATGCTGCA	1045
Qy	968	GCACCGGCCCATGCTGAGAGAGAGATGCTGGAGGCTGTTCAGAAACTTAAATGAAGATGC	1027
Db	1046	GCACCGGCCCATGCTGAGAGAGAGATGCTGGAGGCTGTTCAGAAACTTAAATGAAGATGC	1105
Qy	1028	CTTCTGTGTGAGAGCGGACCCCTGCACTGCGCCATGCACCCGGACCGGCCCTTACTCATCA	1087
Db	1106	CTTCTGTGTGAGAGCGGACCCCTGCACTGCGCCATGCACCCGGACCGGCCCTTACTCATCA	1165
Qy	1088	GACTGGTGTCCAGTTTACCACGAAATCAGGTGTGCTGCTCAAAATTTCTGATGTTGAATTA	1147
Db	1166	GACTGGTGTCCAGTTTACCACGAAATCAGGTGTGCTGCTCAAAATTTCTGATGTTGAATTA	1225
Qy	1148	TCAGCTTAAAAATTTAAAGTGCATTATATAAGACTCTGGGGATGTTGCTGCTCCCTCAAGG	1207
Db	1226	TCAGCTTAAAAATTTAAAGTGCATTATATAAGACTCTGGGGATGTTGCTGCTCCCTCAAGG	1285
Qy	1208	GTCCTCGAAATTTAAACATTTCTGGGCGACGAACAAAAAGTATACATGAGAGAGCTTAA	1267
Db	1286	GTCCTCGAAATTTAAACATTTCTGGGCGACGAACAAAAAGTATACATGAGAGAGCTTAA	1345
Qy	1268	CAAGCGCAGCCTGTCTGCGAGAGTTCAAGCACCTGACCCCTTAGGGAGCAGAGATGTGGAA	1327
Db	1346	CAAGCGCAGCCTGTCTGCGAGAGTTCAAGCACCTGACCCCTTAGGGAGCAGAGATGTGGAA	1405
Qy	1328	TGGAGCGCCTGCCAATTTGTGATCCCTCTTGATGCTGACTGAGAGAGCTGCACCTGATCAC	1387
Db	1406	TGGAGCGCCTGCCAATTTGTGATCCCTCTTGATGCTGACTGAGAGAGCTGCACCTGATCAC	1465
Qy	1388	CTTTCGAGACTGAGGTGTACCAACCAAGGCTCCAAAGTTGACCTAGAGACCACCTCCTTGCC	1447
Db	1466	CTTTCGAGACTGAGGTGTACCAACCAAGGCTCCAAAGTTGACCTAGAGACCACCTCCTTGCC	1525
Qy	1448	AGTTTGTGTGATCTCCACAATCTGTGTGATGCGCCAAATGCTTGGGCATCAATCTGTGTGA	1507
Db	1526	AGTTTGTGTGATCTCCACAATCTGTGTGATGCGCCAAATGCTTGGGCATCAATCTGTGTGA	1585
Qy	1508	TAAACATGCTGACCAATTAACCCCCAAGAACGTGAACCTTTCACTAAGCCGCAATTTGGAAC	1567
Db	1586	TAAACATGCTGACCAATTAACCCCCAAGAACGTGAACCTTTCACTAAGCCGCAATTTGGAAC	1645
Qy	1568	CTGGGACCAAGTGGCGAGGTGCTCACTGACTGGCAATTCCTGTCACCAACCAAGCAGGGCT	1627
Db	1646	CTGGGACCAAGTGGCGAGGTGCTCACTGACTGGCAATTCCTGTCACCAACCAAGCAGGGCT	1705
Qy	1628	GACCATTCGAGCAGCTGACCAACGCTGGCTGAGAAAGCTCCTAGGCCCTGTGTGAACCTACTC	1687
Db	1706	GACCATTCGAGCAGCTGACCAACGCTGGCTGAGAAAGCTCCTAGGCCCTGTGTGAACCTACTC	1765
Qy	1688	AGGGTGTCAGATCATCATGGGCTAAATTTCTGCAAGAAACAAATGCGCTGGCAAGGGCTTCTC	1747
Db	1766	AGGGTGTCAGATCATCATGGGCTAAATTTCTGCAAGAAACAAATGCGCTGGCAAGGGCTTCTC	1825
Qy	1748	CTTCTGGGCTGCTAGACAAATATCATCGACTCTTGTGAAAAAGATATCTTGGCCCTTGG	1807
Db	1826	CTTCTGGGCTGCTAGACAAATATCATCGACTCTTGTGAAAAAGATATCTTGGCCCTTGG	1885

Qy	1808	GAATGAAGGTACATCATATGGGTTTCAATCAGCAAGAGACCGGAGCGGCCATCTTAAGCAC	1867
Db	1886	GAATGAAGGTACATCATATGGTTTTCATACCAAGAGCGCGGAGCGGCATCTTAAGCAC	1945
Qy	1868	AAAGCCCCCGGCGACCTTCTACAGCGCTTACAGAGACACCAAAAGAGAGCGGTAC	1927
Db	1946	AAAGCCCCCGGCGACCTTCTACAGCGCTTACAGAGACAGCAAAAGAGAGGGGTAC	2005
Qy	1928	TTTTCATTGGGTGGAAAAAGACATCACTGTGCAAGACCCAGATCCAGTCTGTAGACCTA	1987
Db	2006	TTTTCACTGGGTGGAAAAAGACATCACTGTGCAAGACCCAGATCCAGTCTGTAGACCTA	2065
Qy	1988	CACCAAGACAGACGCTGAACAACATGTATTTGCTGAATATCATATGGGCTATTAAGTAT	2047
Db	2066	CACCAAGACAGACGCTGAACAACATGTATTTGCTGAATATCATATGGGCTATTAAGTAT	2125
Qy	2048	GGATCGAGCAACATCCGTGGTGTCTCCACTTGTCTACCTCTACCCCGACATTTCCAAAGA	2107
Db	2126	GGATCGAGCAACATCCGTGGTGTCTCCACTTGTCTACCTCTACCCCGACATTTCCAAAGA	2185
Qy	2108	GGAGGCAATTTGGAAAGTACTGTAGGCCGAGAGACCCAGAGACACCCGAGCGCACCG	2167
Db	2186	GGAGGCAATTTGGAAAGTACTGTAGGCCGAGAGACCCAGAGACACCCGAGCGCACCG	2244
Qy	2168	TGAGTCTGCCCCGTACTCTGAAGACCAAGTTCTATCTGTGTACACCAAGACCTGCAGCA	2227
Db	2245	--GTCTGCCCCGTACTCTGAAGACCAAGTTCTATCTGTGTACACCAAGACCTGCAGCA	2302
Qy	2228	TACCATTTAGCTCGCCGATATGTCCCCCGCACCTTTAGATCATATGATATGCAATTA	2287
Db	2303	TACCATTTAGCTCGCCGATATGTCCCCCGCACCTTTAGATCATATGATATGCAATTA	2366
Qy	2288	CGGTGAAGGTGTGAGGCCCTCAGCAGAGAGGGCAGTTGAGTCCGCTCAGCTTTGACATGA	2347
Db	2363	CGGTGAAGGTGTGAGGCCCTCAGCAGAGAGGGCAGTTGAGTCCGCTCAGCTTTGACATGA	2422
Qy	2348	TCTGACCTTGGAGTGTGCTACCTCCCCCATGTGTAGAGAGCTGAACCAAGAGCTGCAGAGA	2407
Db	2423	TCTGACCTTGGAGTGTGCTACCTCCCCCATGTGTAGAGAGCTGAACCAAGAGCTGCAGAGA	2483
Qy	2408	CGGACCTTAGACACCTGCCCCCGGTGCACACCCCTAAGCAGCCGACCCATATCGCTG	2467
Db	2483	CGGACCTTAGACACCTGCCCCCGGTGCACACCCCTAAGCAGCCGACCCATATCGCTG	2544
Qy	2468	AAACTCTCACTTTGTGTGCTCAGA--TTTTTTTTTTAAATTTCCCTACTTCTGCTATCTT	2526
Db	2543	AAACTCTCACTTTGTGTGCTCAGATTTTTTTTTTTAAATTTCTACTTCTGCTATCTT	2603
Qy	2527	GGGCAATCTGGGCACTTTTAAAGAGAGAAATAGTAGAGTGGGGGATTAACCTTTAT	2586
Db	2603	GGGCAATCTGGGCACTTTTAAAGAGAGAAATAGTAGAGTGGGGGATTAACCTTTAT	2666
Qy	2587	GTAAAGAGAGAG--ACCTCTGAGTCTGGGATGGGGCTGAGACAGAGGAGGCAAGG	2641
Db	2663	GTAAAGAGAGAGACACCTCTGTGATCTGGGATGGGGCTGAGAGACAGAGGAGG--CTAAGG	2720
Qy	2646	GGAACACCTCTCTGTCTGTCCGCGCTCGCCCTCCTTTTCACAGACTCGGGGGTGTGT	2709
Db	2721	GGAACACCTCTCTGTCTGTCCGCGCTCGCCCTCCTTTTCACAGACTC--GGGGTGTGTGT	2777
Qy	2706	AGACAGTAGTCTCTGTGTGCCATGGCTAGCTGTATGGCCCACTGTGACTATGATACCC	2766
Db	2780	AGACAGTAGTCTCTGTGTGCCAT--GCATCTGTATGGCCCACTGTGACTATGATACCC	2833
Qy	2766	ATTCTGGGAATCTGTGCTGTGACATTTCAACTTGTCTAATATTCACATAGAAAGCTAGGA	2822
Db	2839	AGGCT--GGAACCTCTGGGCTCTGCACTTTCACACTTGTCTAATATTCACATAGAAAGCTAGGA	2893
Qy	2826	CTAAGCCCAAGAGGTCTCTTTAAAT	2853
Db	2898	CTAAGCCCA--GAGGTCTCTTTAAAT	2924

QY	1811	TGAAGGTACATCATGGTTCATCATCAGCAGAGGAGGAGGCGCCATCTCTTAAGCACA	1870
Db	2001	TGAAGGTACATCATGGTTCATCATCAGCAGAGGAGGAGGCGCCATCTCTTAAGCACA	2060
QY	1871	GGCCCCGGGACCTTCCTACTGCGCTTCAGGAGAGCAGCAAAAGAGAGGGGTCACTTT	1930
Db	2061	GGCCCCGGGACCTTCCTACTGCGCTTCAGGAGAGCAGCAAAAGAGAGGGGTCACTTT	2120
QY	1931	CACTTGGGTGAAAAAGACATCTAGTGGCAGAGCCAGATCCAGTCTGTAGAGCATAC	1990
Db	2121	CACTTGGGTGAAAAAGACATCTAGTGGCAGAGCCAGATCCAGTCTGTAGAGCATAC	2180
QY	1991	CAGAGGACGCTGAAACAACATGTCTTTTGCGAAATCTATCTGCGCTTAAGATCATGA	2050
Db	2181	CAGAGGACGCTGAAACAACATGTCTTTTGCGAAATCTATCTGCGCTTAAGATCATGA	2240
QY	2051	TGCGACCAACATCTGCTGTCTCTCAGCTTGTACTTACCTTACCCGCACTTCCCAAGAGA	2110
Db	2241	TGCGACCAACATCTGCTGTCTCTCAGCTTGTACTTACCTTACCCGCACTTCCCAAGAGA	2300
QY	2111	GGCATTTGGAAGTACTGTAGAGCCCGAGAGCCAGAGACCCCGAAGCCGACAGTAG	2170
Db	2301	GGCATTTGGAAGTACTGTAGAGCCCGAGAGCCAGAGACCCCGAAGCCGACAGTAG	2360
QY	2171	TGCTGCCCCGTACTCTGAAGACCAAGTTACTGTGTGACACCAACGACCTCAGCAATAC	2230
Db	2361	TGCTGCCCCGTACTCTGAAGACCAAGTTACTGTGTGACACCAACGACCTCAGCAATAC	2420
QY	2231	CATTACCTTCGCGATGTCCCCCGCAGCTTAAATTCATTTAGTGCAGTTTGGAAATACGG	2290
Db	2421	CATTACCTTCGCGATGTCCCCCGCAGCTTAAATTCATTTAGTGCAGTTTGGAAATACGG	2480
QY	2291	TGAAGTGCCTGAGCCCTCAGAGAGAGGAGATTAGTGCCTCAGCTTGGACATGATCT	2350
Db	2481	TGAAGTGCCTGAGCCCTCAGAGAGAGGAGATTAGTGCCTCAGCTTGGACATGATCT	2540
QY	2351	GACCTGAGAGTGTCTACTCTCCCATGTGAGAGAGCTGAAACCAAGCTGTAGAGAGCT	2410
Db	2541	GACCTGAGAGTGTCTACTCTCCCATGTGAGAGAGCTGAAACCAAGCTGTAGAGAGCT	2600
QY	2411	GACCTGAGAGACTGCCCCGTCTCCACCCCTAAGCAGCGCAACCCCATATC	2462
Db	2601	GACCTGAGAGACTGCCCCGTCTCCACCCCTAAGCAGCGCAACCCCATATC	2652
RESULT	10		
LOCUS	188784	2652 bp	DNA
DEFINITION	Sequence 8 from patent US 5719042.		PAT
ACCESSION	188784		
VERSION	188784.1	GI:3408724	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2652)		
AUTHORS	Kishimoto, T. and Akira, S.		
TITLE	Nucleic acids encoding transcription factor APXf (acute phase response factor)		
JOURNAL	Patent: US 5719042-A 8 17-FEB-1998;		
FEATURES	Location/Qualifiers		
source	1..2652		
BASE COUNT	683 a	706 c	748 g
ORIGIN		515 t	
Query Match	84.2%;	Score 2415.6;	DB 6; Length 2652;
Best Local Similarity	99.3%;	Pred. No. 0;	
Matches 2436;	Conservative	0; Mismatches	14; Indels
			2; Gaps
QY	13	CAGGCGGCGCAGTGGGC--TCAGCCCGGAGACAGTCGAGACCCCTGACTGCAGCAGAT	70

Db	201	CAGCGCCGGCAGTGGGGCTCAGCCCCGGAGACAGTCCAGACCCCTGACTGTCAGACAGAT	260
Qy	71	GGCTCAGTGGAAACACAGCTGCAGCAGCTGCAGACACAGCTTACCTGAAGCAGCTGCACAGCT	130
Db	261	GGCTCAGTGGAAACACAGCTGCAGCAGCTGCAGACACAGCTTACCTGAGAGCAGCTGCACAGCT	320
Qy	131	GTACAGGACACAGTTCCTCCATGAGACTGGCGCAGTTCTGGACCTTGGATGAGATCA	190
Db	321	GTACAGGACACAGTTCCTCCATGAGAGTGGCGCAGTTCTGGACCTTGGATGAGATCA	380
Qy	191	AGACGGGCATATGCAGCCAGCCAAAGAGTACATGTCACAGTGTGTTTCAATCTCTT	250
Db	381	AGACGGGCATATGCAGCCAGCCAAAGAGTACATGTCACAGTGTGTTTCAATCTCTT	440
Qy	251	GGGTGAAATTCACAGCATATATGCCGATTCCTGCAAGAGTGCATGTCTCTATCAGCA	310
Db	441	GGGTGAAATTCACAGCATATATGCCGATTCCTGCAAGAGTGCATGTCTCTATCAGCA	500
Qy	311	CACCTTGGAAATTCACAGCATATATGCCGATTCCTGCAAGAGTGCATGTCTCTATCAGCA	370
Db	501	CACCTTGGAAATTCACAGCATATATGCCGATTCCTGCAAGAGTGCATGTCTCTATCAGCA	560
Qy	371	TGCCCGGATCCTGGCCCGGATGCTGTGGAAGTCTGCCCTCTCCAGACGGCAGCAC	430
Db	561	TGCCCGGATCCTGGCCCGGATGCTGTGGAAGTCTGCCCTCTCCAGACGGCAGCAC	620
Qy	431	GGCACCACGCAAGGGGGCCAGGCAACACCCCAACAGCCGGCGTAGTACAGAGAACA	490
Db	621	GGCACCACGCAAGGGGGCCAGGCAACACCCCAACAGCTGCCGTAGTACAGAGAACA	680
Qy	491	GCAGATGTTGGAGCAGCATCTTCAGATGTCCGGAAGCGAGTGCAGAGATCTAGAACAGAA	550
Db	681	GCAGATGTTGGAGCAGCATCTTCAGATGTCCGGAAGCGAGTGCAGAGATCTAGAACAGAA	740
Qy	551	AATGAAGGTGGGGAAGAACCTCCAGAGCAGCTTGATTTCACTACAAAACCTCAACAG	610
Db	741	AATGAAGGTGGGGAAGAACCTCCAGAGCAGCTTGATTTCACTACAAAACCTCAACAG	800
Qy	611	CCAGAGACATATGCAGATCTGAAATGGAACCAACCAAGTGTGTACCCAGACAGACATGCA	670
Db	801	CCAGAGACATATGCAGATCTGAAATGGAACCAACCAAGTGTGTACCCAGACAGACATGCA	860
Qy	671	GCAGCTGGAACAGATGCTACAGCCCTGCAGACAGATGCCGGAAGAGATGTGAGTGAAGCT	730
Db	861	GCAGCTGGAACAGATGCTACAGCCCTGCAGACAGATGCCGGAAGAGATGTGAGTGAAGCT	920
Qy	731	GGCGGGGCTTGTGTAGCAATGGAAGTACGTGCAGAAAGACACTGACTGATGAAGAGCTGGC	790
Db	921	GGCGGGGCTTGTGTAGCAATGGAAGTACGTGCAGAAAGACACTGACTGATGAAGAGCTGGC	980
Qy	791	TGCACTGGAAGAGGGCGCCAGAGATTCGGTGCATGGAAGGCCCTCCCAACATGTGCTGGA	850
Db	981	TGCACTGGAAGAGGGCGCCAGAGATTCGGTGCATGGAAGGCCCTCCCAACATGTGCTGGA	1040
Qy	851	CCGCTGGAAGAACTGGATTAACCTATTAGCAGATCTCAACTTCAGACCCCGCCACAAGAT	910
Db	1041	CCGCTGGAAGAACTGGATTAACCTATTAGCAGATCTCAACTTCAGACCCCGCCACAAGAT	1100
Qy	911	TAAGAACCTGGAGAGCTGCAGACAGAAAGTGTCTTACAAAGGGCAGCCCTATGTGCACAA	970
Db	1101	TAAGAACCTGGAGAGCTGCAGACAGAAAGTGTCTTACAAAGGGCAGCCCTATGTGCACAA	1160
Qy	971	CCGGCCCATATGCTGGAGAGAGAGATGTGTGAGACTTTCAGAAACTTAATCAAGAGCTCT	1030
Db	1161	CCGGCCCATATGCTGGAGAGAGAGATGTGTGAGACTTTCAGAAACTTAATCAAGAGCTCT	1220
Qy	1031	CGTGTGAGAGCGGACGCGCTGCATGCCCATGCACACCGGACCGGCTTATGTCATCAAGAC	1090
Db	1221	CGTGTGAGAGCGGACGCGCTGCATGCCCATGCACACCGGACCGGCTTATGTCATCAAGAC	1280
Qy	1091	TGCTGTCCAGTTTACCACAGAAAGTCAAGTTCTGTCTGCTCAAAATTCCTGAGTGAATTTACA	1150
Db	1281	TGCTGTCCAGTTTACCACAGAAAGTCAAGTTCTGTCTGCTCAAAATTCCTGAGTGAATTTACA	1340

QY	1151	GCTTAAATTAAGTGTGCATTGATTAAGACACTCTGGGGAATTTGCTGCTCCACAGAGGCTC	1211
Db	1341	GCTTAAATTAAGTGTGCATTGATTAAGACACTCTGGGGAATTTGCTGCTCCACAGAGGCTC	1400
QY	1211	TCGGAATTTAAACATTCTGGGCGACGAAACCAAAAGTGTGAACATGAGAGTCTAACAA	1270
Db	1401	TCGGAATTTAAACATTCTGGGCGACGAAACCAAAAGTGTGAACATGAGAGTCTAACAA	1460
QY	1271	CGGAGCGCTTCTCGACAGATTGAAGACCTGACCTTGGAGGACGAGATGTGGGAATGG	1330
Db	1461	CGGAGCGCTTCTCGACAGATTGAAGACCTGACCTTGGAGGACGAGATGTGGGAATGG	1520
QY	1331	AGGCGGTGCCAATTTGATGATCCCTCTTGTATCGTGAAGAGGAGCTGACCTGATCACCTT	1390
Db	1521	AGGCGGTGCCAATTTGATGATCCCTCTTGTATCGTGAAGAGGAGCTGACCTGATCACCTT	1580
QY	1391	CGAAGCTGAGGTGTACCAACCAAGCGCTCAAGATTTACCTAGAGACCCACCTCTTGCCAGT	1450
Db	1581	CGAAGCTGAGGTGTACCAACCAAGCGCTCAAGATTTACCTAGAGACCCACCTCTTGCCAGT	1640
QY	1451	TGTGTGATCTCCAAACATCTGTGCAGATGCCAAATCTTGGGCATCAATCTGTGTATAA	1510
Db	1641	TGTGTGATCTCCAAACATCTGTGCAGATGCCAAATCTTGGGCATCAATCTGTGTATAA	1700
QY	1511	CATGCTGACCAATTAACCCCAAGAAAGCTGAAAGTCTTCACTAAGCCGCCAATTTGGAACCTG	1570
Db	1701	CATGCTGACCAATTAACCCCAAGAAAGCTGAAAGTCTTCACTAAGCCGCCAATTTGGAACCTG	1760
QY	1571	GGACCAAGTGGCGGAGGTGCTCAGCTGGGCAATTCTGTGCACACCAACGACGAGGCTGAG	1630
Db	1761	GGACCAAGTGGCGGAGGTGCTCAGCTGGGCAATTCTGTGCACACCAACGAGGCTGAG	1820
QY	1631	CATGAGACAGGTGACCAACGCTGGCTAGAAAGCTCCTAGGGGCTGTGTGAACACTACTAGG	1690
Db	1821	CATGAGACAGGTGACCAACGCTGGCTAGAAAGCTCCTAGGGGCTGTGTGAACACTACTAGG	1880
QY	1691	GTGTGATGATCATGATGGGCTAAATTTCTGCAGAAAGAAACATGCTGGCCAAAGGCTTCTCCTT	1750
Db	1881	GTGTGATGATCATGATGGGCTAAATTTCTGCAGAAAGAAACATGCTGGCCAAAGGCTTCTCCTT	1940
QY	1751	CTGGGTCTGGGTATACCAATTCATTCGACACCTTGTGAAAAGTATATCTTGGCCCTTGGAA	1810
Db	1941	CTGGGTCTGGGTATACCAATTCATTCGACACCTTGTGAAAAGTATATCTTGGCCCTTGGAA	2000
QY	1811	TGAAGGTCATCATGCGGTTTCATCAGCAAGACGCGGAGCGGCCATCTTAAGCACAA	1870
Db	2001	TGAAGGTCATCATGCGGTTTCATCAGCAAGACGCGGAGCGGCCATCTTAAGCACAA	2060
QY	1871	GCCCCCGGGACCTTCTCTACTGCGGCTTCACGCGACGACCAAGAAAGAGGGGTCACTTT	1930
Db	2061	GCCCCCGGGACCTTCTCTACTGCGGCTTCACGCGACGACCAAGAAAGAGGGGTCACTTT	2120
QY	1931	CACCTTGGTGGAAAAAGACATCACTGCGCAAGACCCAGATCACTCTGTAGAGCATACAC	1990
Db	2121	CACCTTGGTGGAAAAAGACATCACTGCGCAAGACCCAGATCACTCTGTAGAGCATACAC	2180
QY	1991	CAGACGAGGTGAAACAACATATGTCATTGTGTGAATTCATCATGAGGCTATTAAGATCATGCA	2050
Db	2181	CAGACGAGGTGAAACAACATATGTCATTGTGTGAATTCATCATGAGGCTATTAAGATCATGCA	2240
QY	2051	TTCGACCAACATCTCGTGTCTCAGCTTGTCTACCTCTACCCGACATTCCAAAGGAGGA	2110
Db	2241	TTCGACCAACATCTCGTGTCTCAGCTTGTCTACCTCTACCCGACATTCCAAAGGAGGA	2300
QY	2111	GGCATTTTGAAAGTACTGTATAGGCGCGAGACCCAGAGACACCCGAAAGCGACCCAGGTAG	2170
Db	2301	GGCATTTTGAAAGTACTGTATAGGCGCGAGACCCAGAGACACCCGAAAGCGACCCAGGTAG	2360
QY	2171	TGCTGCCCCGTCACCTGAAAGACCAATTCATCTGTGTGACACCAAGCACTGTGACCAATAC	2230
Db	2361	TGCTGCCCCGTCACCTGAAAGACCAATTCATCTGTGTGACACCAAGCACTGTGACCAATAC	2420

QY	2231	CATTGACTGCCGAGATGTCCTCCCGACACTTTGATTCATTTGATGACGATTTGGAAATTAACGG	2290
Db	2421	CATTGACCTGCGGAGATGTCCTCCCGACACTTTGATTCATTTGATGACGATTTGGAAATTAACGG	2480
QY	2291	TGAAGGTCCTAGAGCCCTCAGAGAGAGGCGACATTTGAGTCAGCTTGTGACATGATCT	2350
Db	2481	TGAAGGTCCTAGAGCCCTCAGAGAGAGGCGAGTTTGATGTCGCTACGTTTGCATGATCT	2540
QY	2351	GACCTCGAGCTGTGCTACCTCCCGCATGTGAGAGACTGAACCAAGAACTGCAGAGACT	2410
Db	2541	GACCTCGAGCTGTGCTACCTCCCGCATGTGAGAGACTGAACCAAGAACTGCAGAGACT	2600
QY	2411	GACTGAGACACCTGCGCTGCTCACCCTCAACGACCGGACCCCATATC	2462
Db	2601	GACTGAGACACCTGCGCTGCTCACCCTCAACGACCGGACCCCATATC	2652
RESULT	11		
MUSAPRF			
LOCUS	MUSAPRF	2652 bp	mRNA
DEFINITION	Mus musculus acute phase response factor (APRF) gene, complete cds.	ROD	02-MAY-1994
ACCESSION	L29278		
VERSION	L29278.1	GI:476715	
KEYWORDS	DNA-binding protein; acute phase response factor; transcription factor.		
SOURCE	Mus musculus (strain BALB/c, sub-species domesticus) (library: lambda gtl1 cDNA library) liver cDNA to mRNA.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2652)		
AUTHORS	Akira,S., Nishio,Y., Inoue,M., Wang,X.-J., Shi,W., Matsusaka,T., Yoshida,K., Sudo,T., Naruto,M., and Kishimoto,T.		
TITLE	Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway		
JOURNAL	Cell 77, 63-71 (1994)		
MEDLINE	94208062		
FEATURES			
SOURCE	Location/Qualifiers		
	1..2652		
	/organism="Mus musculus"		
	/strain="BALB/c"		
	/sub_species="domesticus"		
	/db_xref="taxon:10090"		
	/tissue_type="liver"		
	/tissue_id="lambda gtl1 cDNA library"		
	259..2571		
	/gene="APRF"		
	259..2571		
	/gene="APRF"		
	/note="ISGF3 p91-related transcription factor"		
	/codon_start=1		
	/function="transcription factor involved in the gp130-mediated signaling pathway"		
	/product="acute phase response factor"		
	/protein_id="AAA37254.1"		
	/db_xref="GI:476716"		
	/translation="MAONNQLOOLDTRLLEOLHQLYSQSFPEMLROPLAPMIESDWMYAKASRSHATLVFNLGELIDQYSRFLQESVNTLOHMLRIRKOPLOSRYLEKPELIRKVARLWEESRLDLOATAAQGGQAMHPVLAATVEQOMLEHLDVRRVYDLEGRKMKVAENLDQDFNYKTLKSGDDMLNGNSVTRQKMOOLEQMLTALDOMRSIYQSLSELALISMEYVOKTIDELDALDKRRQIACIGCPNIDCLRLNNITSLAESOLDLITRQOIKKLLELOOKKYSYKGPVIOHRRPILBEELVLFNLKSAFVVRQCMHPHPIRPLVKTGVQPTTKRLRLVYKPELNYQLIKYCIDKSGDVAALAGSKFNILGNTNVTNDESSNNGSLSAEFKHLTLREORCGNGRANCASLIVBELHLITFEVYHOGIKIDTESLSPVYVINSICOMPNAWASILMYINMLTNPKNVNVEFKPIGTVDQVAEVIKMEQSTTKRGSLISOLTTLAKLEKLGPVNSGCOITWAKCKENMAGKGSFWMLNIDLVKRYTILALNMGYIMGFIISKERPAVILSTPQGTFLRSESSEKGVFTWMLKSDISGKQIOISVEPRTSKAOLNMFALIMYGKIMDTNIIIVSLYLYDGIPEEZEFGKYRPPSOEHPADSGSAAPMYLTKTKTCVPTTCSNTIDLPNSLDMQFNNNGGGAPEPSAGGQESLTFMDLTLSEATSPM"		
BASE COUNT	683 a	706 c	748 g
ORIGIN			515 t

OY 2111 GGCATTGGAAGTACTGTAGGCCGAGAGACCCGGAAGCCGACCCAGCTAG 2170
|||||
DB 2301 GGCATTGGAAGTACTGTAGGCCGAGAGACCCGGAAGCCGACCCAGCTAG 2160
OY 2171 TGCTGCCCCGCTGAGAGCCAGTTCCTGTGTGACACCAAGACCTGGCGCAATAC 2230
|||||
DB 2361 TGCTGCCCCGCTGAGAGCCAGTTCCTGTGTGACACCAAGACCTGGCGCAATAC 2420
OY 2231 CATTGACCTGCGGATGTCGCCCGGACCTTATGATTCAGTTGGAAATTAACGG 2290
|||||
DB 2421 CATTGACCTGCGGATGTCGCCCGGACCTTATGATTCAGTTGGAAATTAACGG 2480
OY 2291 TGAAGTGTGTGAGCCCTCAGAGAGAGAGGAGTTGAGTGGCTACGTTTGACATGATCT 2350
|||||
DB 2481 TGAAGTGTGTGAGCCCTCAGAGAGAGAGGAGTTGAGTGGCTACGTTTGACATGATCT 2540
OY 2351 GACCTGCGAGTGTGTACTCCCGCATGTGAGAGCTGAAACAGAGAGCTGAGAGAGCT 2410
|||||
DB 2541 GACCTGCGAGTGTGTACTCCCGCATGTGAGAGCTGAAACAGAGAGCTGAGAGAGCT 2600
OY 2411 GACTTGAGACACCTGCGCCGCTGCTCCACCCCTTAAGCAGCCGGAACCCCATATC 2462
|||||
DB 2601 GACTTGAGACACCTGCGCCGCTGCTCCACCCCTTAAGCAGCCGGAACCCCATATC 2652

RESULT 12
E10408 standard; RNA; ROD; 2652 BP.
XX E10408:
XX
SV E10408.1
XX
XX 08-OCT-1997 (Rel. 52, Created)
DT 02-SEP-2000 (Rel. 65, Last updated, Version 2)
XX
XX cDNA encoding an acute phase response factor.
DE
XX JP 1995324096-A/2.

OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
XX
XX [1]
XX 1-2652
XX Shunliyou S., Kishimoto C.:
XX "TRANSCRIPT APRF".
XX Patent number JP1995324096-A/2, 12-DEC-1995.
XX KISHIMOTO CHUZO.

XX
XX OS Mus sp. (mouse)
CC PN JP 1995324096-A/2
CC PD 12-DEC-1995
CC PF 03-APR-1995 JP 1995077584
CC PR 04-APR-1994 JP 94P 65825
CC PI SHINRIYOU SHIZUO, KISHIMOTO CHUZO
CC PC C07K14/52,A61K31/70,A61K38/00,A61K39/395,A61K48/00,C07H21/04,
CC PC C12P21/02//C12N1/21,C12N5/10,(C12P21/02,C12R1:91),(C12N1/21,
CC PC C12R1:19),
CC PC (C12N5/10,C12R1:91);
CC CC strandedness: Double;
CC CC topology: Linear;
CC CC Key Location/Qualifiers
CC FH source 1..2652
CC FT /organism="Mus sp."
CC FT /tissue="liver"
CC FT CDS 259..2571
CC FT /product="APRF"
XX
FH Key Location/Qualifiers

FH
FT source 1..2652
FT /db_xref="taxon:10095"
FT /organism="Mus sp."
XX
SQ Sequence 2652 BP; 683 A; 706 C; 748 G; 515 T; 0 other;

Query Match 84.2%; Score 2415.6; DB 22; Length 2652;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

OY 13 CAGCCGCGCAGTCTGGGC--TCAGCCCGGAGACAGTGTGAGACCCCTGACTGCGAGAT 70
|||||
DB 201 CAGCCGCGCAGTCTGGGCCTCAGCCCGGAGACAGTGTGAGACCCCTGACTGCGAGAT 260
OY 71 GGCCTAGTGAACACCTGCGACAGCTGCGACACAGCTTACCTGAGCAGCTGACAGCT 130
|||||
DB 261 GGCCTAGTGAACACCTGCGACAGCTGCGACACAGCTTACCTGAGCAGCTGACAGCT 320
OY 131 GTACAGCGACACCTTCCCATGAGAGCTCGCGCAGTTCTGGCACCTTGGATTGAGAGTCA 190
|||||
DB 321 GTACAGCGACAGCTTCCCATGAGAGCTCGCGCAGTTCTGGCACCTTGGATTGAGAGTCA 380
OY 191 AGACTGGGATATGACAGCCGCAAGAGTGCATGCGCAGCTGCTGCTGCTGCTGCTGCT 250
|||||
DB 381 AGACTGGGATATGACAGCCGCAAGAGTGCATGCGCAGCTGCTGCTGCTGCTGCTGCT 440
OY 251 GGGTGAATTTGACACCAATATAGCCGATTCCTGCAAGAGTCCCAATGCTCTATGAGCA 310
|||||
DB 441 GGGTGAATTTGACACCAATATAGCCGATTCCTGCAAGAGTCCCAATGCTCTATGAGCA 500
OY 311 CAACCTTCGAGAAATCAAGAGAGTTCTGACAGACAGTATCTTGAGAACCAATGGAAT 370
|||||
DB 501 CAACCTTCGAGAAATCAAGAGAGTTCTGACAGACAGTATCTTGAGAACCAATGGAAT 560
OY 371 TGCCCGGATGCTGGCCGATGCTGTGGAGAGTCTCGCTCTCTCCAGACGGCAGCCAC 430
|||||
DB 561 TGCCCGGATGCTGGCCGATGCTGTGGAGAGTCTCGCTCTCTCCAGACGGCAGCCAC 620
OY 431 GCGAGCCCGCAAGGGGGCCAGGCCCAACCAACCCAGCCGCTGATGACAGAGCA 490
|||||
DB 621 GCGAGCCCGCAAGGGGGCCAGGCCCAACCAACCCAGCCGCTGATGACAGAGCA 680
OY 491 GCGAGTGTGAGCAGCAGTCTTCCAGATGTCGGAAGCGAGTCCAGATTTAGACAGAA 550
|||||
DB 681 GCGAGTGTGAGCAGCAGTCTTCCAGATGTCGGAAGCGAGTCCAGATTTAGACAGAA 740
OY 551 AATGAAGGTGAGAGAACCTCCAGAGCAGCTTGTGATTTCACTAGCAAAACCTCAAGAG 610
|||||
DB 741 AATGAAGGTGAGAGAACCTCCAGAGCAGCTTGTGATTTCACTAGCAAAACCTCAAGAG 800
OY 611 CCAAGAGACATGACAGATCTGAAATGAAACCAACCACTGTGTACCAAGAGAAATGCA 670
|||||
DB 801 CCAAGAGACATGACAGATCTGAAATGAAACCAACCACTGTGTGTACCAAGAGAAATGCA 860
OY 671 GCACTGTGAACAGATCTCAGACAGCCCTGACACAGATGCGGAGAGAGATTGTGAGTAC 730
|||||
DB 861 GCACTGTGAACAGATCTCAGACAGCCCTGACACAGATGCGGAGAGAGATTGTGAGTAC 920
OY 731 GCGGGGCTCTTGTGACGAATGAGTACGTGCAAGAACACTGATGATGAAGAGCTGCG 790
|||||
DB 921 GCGGGGCTCTTGTGACGAATGAGTACGTGCAAGAACACTGATGATGAAGAGCTGCG 980
OY 791 TGACTGGAAGAGGCGGCGAGAGATGCGGTGCAATGAGAGGCGCTCCCAACATCGCTGGA 850
|||||
DB 981 TGACTGGAAGAGGCGGCGAGAGATGCGGTGCAATGAGAGGCGCTCCCAACATCGCTGGA 1040
OY 851 CCGTCTGGAAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 910
|||||
DB 1041 CCGTCTGGAAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1100
OY 911 TAAAGA/ACTGGAAGAGCTGACAGCAAGAGTGTCTTCAAGAGGCGGACCTTATGTGACAGA 970


```

|||||
Db 1101 TAAGAAATCTGAGAGCTGAGCAGCAAAAGTGTCTTACAAAGGCCAACCTATCGTCGACGA 1160
Qy 971 CCGGCCCATGCTGAGAGAGAGATGTCGACCTGTCAGAACTTAATGAAGAGTCCTT 1030
Db 1161 CCGGCCCATGCTGAGAGAGAGATGTCGACCTGTCAGAACTTAATGAAGAGTCCTT 1220
Qy 1031 CGTGTGAGAGGCGGACCCCTGATGCCATGACCCCGAGCGCCCTTAGTCATCAAGAC 1090
Db 1221 CGTGTGAGAGGCGGACCCCTGATGCCATGACCCCGAGCGCCCTTAGTCATCAAGAC 1280
Qy 1091 TGTGTCCAGTTTACCAGAAAGTCAAGTGTCTGTCAAAATTTCTCGAGTTGAATATCA 1150
Db 1281 TGTGTCCAGTTTACCAGAAAGTCAAGTGTCTGTCAAAATTTCTCGAGTTGAATATCA 1340
Qy 1151 GCTTAAATTAAGTGTGATTAAGACCTGTGGGATGTGTGCGCTCAGAGGGTC 1210
Db 1341 GCTTAAATTAAGTGTGATTAAGACCTGTGGGATGTGTGCGCTCAGAGGGTC 1400
Qy 1211 TGGAAATTTAATCTTGGGACAGACAAAGCTGTGATGAACATGAGAGACTTAACAA 1270
Db 1401 TCGGAAATTTAATCTTGGGACAGACAAAGCTGTGATGAACATGAGAGACTTAACAA 1460
Qy 1271 CCGCAGCCTGTCTGAGAGATTAAGACCTGACCTTAGGGAGCAGAGATGTGGAAATGG 1330
Db 1461 CCGCAGCCTGTCTGAGAGATTAAGACCTGACCTTAGGGAGCAGAGATGTGGAAATGG 1520
Qy 1331 AGGCCCTGGCAATTTGATGATGCTGCTTGTATCGGATGAGAGCTGACCTGATCCTT 1390
Db 1521 AGGCCCTGGCAATTTGATGATGCTGCTTGTATCGGATGAGAGCTGACCTGATCCTT 1580
Qy 1391 CGAGACTGAGAGTGTACACCAAGGCTCTCAAGATTGACCTAGAGACCCACTCTTGGCACT 1450
Db 1581 CGAGACTGAGAGTGTACACCAAGGCTCTCAAGATTGACCTAGAGACCCACTCTTGGCACT 1640
Qy 1451 TGTGTGATCTCCAAATCTGTCAAGTGGCAAAATGCTTGGGACATCAATCTGTGATTA 1510
Db 1641 TGTGTGATCTCCAAATCTGTCAAGTGGCAAAATGCTTGGGACATCAATCTGTGATTA 1700
Qy 1511 CATGCTGACCAATTAACCCCAAGAGCTGAACCTTCTTCACTAAGCCGCCAATTTGGAACCTG 1570
Db 1701 CATGCTGACCAATTAACCCCAAGAGCTGAACCTTCTTCACTAAGCCGCCAATTTGGAACCTG 1760
Qy 1571 GGACCAAGTGGCGAGGTGTCTGAGTGGGCTGTCTGCTCCACCAAGGAGGAGGCTGAG 1630
Db 1761 GGACCAAGTGGCGAGGTGTCTGAGTGGGCTGTCTGCTCCACCAAGGAGGAGGCTGAG 1820
Qy 1631 CATGAGCAGCTGACAAAGCTGTGAGAGCTCTAGGGCTGTGTGTAATCTACTCAGG 1690
Db 1821 CATGAGCAGCTGACAAAGCTGTGAGAGCTCTAGGGCTGTGTGTAATCTACTCAGG 1880
Qy 1691 GTGTGACATCATGAGCTAAATTTCTCAAAAGAAACATGCTGGCAAGGGCTTCTCTT 1750
Db 1881 GTGTGACATCATGAGCTAAATTTCTCAAAAGAAACATGCTGGCAAGGGCTTCTCTT 1940
Qy 1751 CTGTGCTGGCTAGCAATATCATGACCTTGTGAAAAAGTATATCTTGGCCCTTTGAA 1810
Db 1941 CTGTGCTGGCTAGCAATATCATGACCTTGTGAAAAAGTATATCTTGGCCCTTTGAA 2000
Qy 1811 TGAAGGATACATCATGAGTTTATCATCAGCAAGAGGAGCGGCGCATCTTAAGCACA 1870
Db 2001 TGAAGGATACATCATGAGTTTATCATCAGCAAGAGGAGCGGCGCATCTTAAGCACA 2060
Qy 1871 GCGCCCGGACCTTCTCTACTAGCGCTTCAAGGAGAGCAGCAAAAGAGAGGCGTCACTT 1930
Db 2061 GCGCCCGGACCTTCTCTACTAGCGCTTCAAGGAGAGCAGCAAAAGAGAGGCGTCACTT 2120
Qy 1931 CACTTGGTGGAAAAAGCATCACTGGCAGAGACCCAGATCACTGTGTAGAGCCATACAC 1990
Db 2121 CACTTGGTGGAAAAAGCATCACTGGCAGAGACCCAGATCACTGTGTAGAGCCATACAC 2180
Qy 1991 CAAGCAGAGCTGAACAACATGTCTATTTGTAATATCATATGGGCTATATAGATCATGGA 2050
|||||

```

```

Db 2181 CAAGCAGCAGCTGAACAACATGTCATTTGCTGGAATATCATGCGGCTATAGATCATGGA 2240
Qy 2051 TGGACCAACATCTCTGTGTCTCCACTTGTCTACCTTACCCGACATTTCCAGAGAGA 2110
Db 2241 TGGACCAACATCTCTGTGTCTCCACTTGTCTACCTTACCCGACATTTCCAGAGAGA 2300
Qy 2111 GGCATTTGAAAGTACTGTAGGCGCCGAGAGCAGACACCCGCAACCCGACAGTAG 2170
Db 2301 GGCATTTGAAAGTACTGTAGGCGCCGAGAGCAGACACCCGCAACCCGACAGTAG 2260
Qy 2171 TGGCGCCCGGACCTGGAAGACCAAGTTCATCTGTGTGACACCAAGCAGCTGACCAATAC 2230
Db 2361 TGGCGCCCGGACCTGGAAGACCAAGTTCATCTGTGTGACACCAAGCAGCTGACCAATAC 2420
Qy 2231 CATTGACCTGGCATGTCTCCCGGACCTTTAGATTCATTTGATGATGCAAGTTTGGAAATTA 2290
Db 2421 CATTGACCTGGCATGTCTCCCGGACCTTTAGATTCATTTGATGATGCAAGTTTGGAAATTA 2480
Qy 2291 TGAAGTGTCTGAGCCCTGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2350
Db 2481 TGAAGTGTCTGAGCCCTGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2540
Qy 2351 GACCTGAGAGTGTGCTTACCTCCCGCATGTGAGAGCTGAACACGAGAGCTGAGAGAGCT 2410
Db 2541 GACCTGAGAGTGTGCTTACCTCCCGCATGTGAGAGCTGAACACGAGAGCTGAGAGAGCT 2600
Qy 2411 GACCTGAGAGCTGCTCCCGGCTGCTCCACCCCTAACCCAGCCGACCTATATC 2462
Db 2601 GACCTGAGAGCTGCTCCCGGCTGCTCCACCCCTAACCCAGCCGACCTATATC 2652

RESULT 13
RNSTAT3PT
LOCUS RNSTAT3PT 2924 bp mRNA ROD 29-FEB-1996
DEFINITION R.norvegicus mRNA for Stat3 protein.
ACCESSION X919.0
VERSION X919.0.1 GI:1107848
KEYWORDS stat3 gene; stat3 protein.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2924)
AUTHORS Ripberger,J.A., Fritz,S., Richter,K., Hocke,G.M., Lotzspeich,F. and
Fey,G.H.
JOURNAL Transcription factors Stat3 and Stat5b are present in rat liver
MEDLINE J. Biol. Chem. 270 (50), 29998-30006 (1995)
TITLE nuclei late in an acute phase response and bind interleukin-6
AUTHORS Ripberger,J.A.
JOURNAL Direct Submission
Submitted (26-SEP-1995) J.A. Ripberger, Universitaet
Erlangen-Nuernberg, Lehrstuhl f. Genetik, Staudtstrasse 5, D 91058
Erlangen, FRG
FEATURES
source
1..2924
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/dev_stage="adult"
/lissue_type="liver, acute phase"
/cell_type="predominantly hepatocytes"
/clone_lib="lambda ZapII (Stratagene)"
1..2313
/gene="Stat3"
1..2313
/gene="Stat3"
1..2313
/gene="Stat3"
/notes="single, polypeptide chain; needs to form homo- or
heterodimers for specific DNA-binding"
/codon_start=1
/product="Stat3 protein"

```

```
/protein_id="CAA62920.1"
/db_xref="GI:1107849"
/translation="MAQMOLOLDTRYLEHOLYSDSPFELROFLAPWIESODMA
VAASKESHATLVFNHNLGEIDDOYSRFLNPNAYVONHRLIKOFLDSRYLEKMEIA
RIVARCEMESRLIOTATAAODGGOAHNEAVVJOHLEOOLDOVHRVYDLEO
KMKVENVLODDFEDFNKTKLSQGMODLNQNSVTKOKNOLEOMITLADOMRSIV
SELGLLSAMEYVOKTILDEBLADMKRRQOIACTIGEPNITCLDLRWMTISLASOLO
TRDOIKRLELOOKVSYKDBIVOHRLMEERIVDLERNLMSAFVVEROPCEMHPD
RPLVKTGVQFTTRKLLVLEPBLNOLKIKVCIDKSDGVLAALSGSRKFNILGTNK
VMNEESNNGSLSAEFKHLTREORCNGGANCDAISLYTEHLITFEYEVHOGIL
KIDLETSLPVLVSNICOMPANASILUMYMLNPNVNFPEKPIGTIDVOYAEVL
SMOESSTKRGSLTEOLITLAEKLGSPGVNSGCOITMAKCFKEMKAGKSPFVMTD
NIIDLVKKYILALMNBGYIMGFIKSEBRALTSKRPPTELLRESSEKBEQVFTFY
EKDISKTQIOSVEPTKQOJNNMSFAEIIINGITIMODTNILVSPLYLIPDIKEEA
FGKYCRPESQHPADPQSAAPYLKTKFICVTPPTCSNTIDLPMSPTLDSLMOFGNN
GEQAPAGQGFESLTFDMDLTSECATSPM"
```

```
BASE COUNT      771 a      732 c      784 g      637 t
ORIGIN
Query Match      80.0%; Score 2296; DB 10; Length 2924;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 2550; Conservative 0; Mismatches 210; Indels 48; Gaps 8;
OY 69 ATGGCTAGTGGACACGCTGCAGCGTGGACACACGCTACCTGAGCAGCTGGACAG 128
Db 1 ATGGCTAGTGGACACGCTGCAGCGTGGACACGCTACCTGAGCAGCTTATCAG 60
OY 129 CTGTACAGCGACACGCTGCCATGAGCGTGGCGAGTTCTGTGACCTTGAATGAG 188
Db 61 CTGTACAGCGAGTCTGCCATGAGCGTGGCGAGTTCTGTGAGTATGAGAGC 120
OY 189 CAAGACTGGCGATATGACCGCCAGCAAGAGTCACATGCCAGCTTGTGTATATCTC 248
Db 121 CAAGATTGGCGATATGACCGCCAGCAAGAGTCACAGCCACTGTGTATATACCTC 180
OY 249 TTGGGTGAATTTGACCCAGCAATATAGCCGATTCCTGCAAGAGTCCCATGCTATCAG 308
Db 181 TTGGGCGAGATGACCGAGATATAGCCGATTCCTGCAAGAGTCCCATGCTATCAG 240
OY 309 CACAACCTTCGAAAGATCAGCAGTTTTCGACAGCAGGATCTTGAGAGCCAAATGGA 368
Db 241 CACAACCTTCGCAAGATCAGCAGTTTTCGACAGCAGGATCTTGAGAGCCAAATGGA 300
OY 369 ATTGCCGAGATCGTGGCCGATGCTGTGGGAAGAGTCTCGCTCCTCCAGAGCGAGCC 428
Db 301 ATTGCCGAGATTTGTGGCCGATGCTGTGGGAAGAGTCTCGCTCCTCCAGAGCGAGCC 360
OY 429 ACGCGAGCCACAGAGGGGGCCAGGCCAACCCACAGACGCGCGCTAGTGAAGAG 488
Db 361 ACGCGAGCCACAGAGGGGGCCAGGCCAACCCACAGACGCTGCTAGTGAAGAG 420
OY 489 CAGCAGATTTGGAGCAGCATCTTGAATGTCGGAGCGAGTCCAGAGATTAACAG 548
Db 421 CAGCAGATTTGGAGCAGCATCTTGAATGTCGGAGCGAGTCCAGAGATTAACAG 480
OY 549 AAAATGAGGTGGTGGAGAACCTCCAGAGCAGCTTTGATTTCACCTACAAAACCTTCAAG 608
Db 481 AAAATGAGGTGGTGGAGAACCTCCAGAGTGCCTTTGATTTCACCTACAAAACCTTCAAG 540
OY 609 AGCCAGAGGAGATCAGAGATTTGAATGGAAACACAGCTCTGTGACCCAGAGAGATG 668
Db 541 AGTCAAGAGGAGATCAGAGATTTGAATGGAAACACAGCTCTGTGACCCAGAGAGATG 600
OY 669 CAGCAGTGGAGACAGATCTCAGACGCTGGACAGATGGGAGAGATTGTGATGAG 728
Db 601 CAGCAGTGGAGACAGATCTCAGACGCTGGACAGATGGGAGAGATTGTGATGAG 660
OY 729 CTGGCGGGGCTCTTGTGAGCAATGAGTACGTGAGAAAGACACTGATGATGAGAGCTG 788
Db 661 CTGGCGGGGCTCTTGTGAGCAATGAGTACGTGAGAAAGACACTGATGATGAGAGCTG 720
```

```
OY 789 CTTGACTGGAAGAGGGCGCCAGAGATTCGCTGATCGGAGGGCCCTCCCAACATCTGCTG 848
Db 721 GCTGACTGGAAGAGGGCGCCAGAGATTCGCTGATCGGAGGGCCCTCCCAACATCTGCTG 780
OY 849 GACCCCTGGAAGAACTGATTAATCTTATACAGAAATCTCAACTTCAGACCCGCAACAA 908
Db 781 GACCCCTGGAAGAACTGATTAATCTTATACAGAAATCTCAACTTCAGACCCGCAACAA 840
OY 909 ATTAAGAACTGAGAGCTGCGAGCAGAAAGTCTCTCAAGGGCGACCTATGCTGAG 966
Db 841 ATTAAGAACTGAGAGCTGCGAGCAGAAAGTCTCTCAAGGGCGACCTATGCTGAG 900
OY 969 CACGGGCCATGCTGGAGAGAGATGCTGGAGCTGTGCAAGAACTTAATAAGAGTCC 1027
Db 901 CACCGTCAATGCTGGAGAGAGATGCTGGAGCTGTGCAAGAACTTAATAAGAGTCC 960
OY 1029 TTGCTGTGAGAGCGGCGACCCCTGATGCCATGCAACCCGAGCCGCTTATGATCAAG 1088
Db 961 TTGCTGTGAGAGCGGCGACCCCTGATGCCATGCAACCCGAGCCGCTTATGATCAAG 1020
OY 1089 ACTGCTGTCCAGTTTACCAAGAACTCAGGTTTCTGTCAAAATTTCTGATGATTAAT 1148
Db 1021 ACTGCTGTCCAGTTTACCAAGAACTCAGGTTTCTGTCAAAATTTCTGATGATTAAT 1080
OY 1149 CAGCTTAAATTAAGTGTGATTAAGACTCTGGGAGATGTTGCTGCTCAGAGAG 1208
Db 1081 CAGCTTAAATTAAGTGTGATTAAGACTCTGGGAGATGTTGCTGCTCAGAGAG 1140
OY 1209 TCTGGAATTTTAACTTGTGGGCGAAGACAAAGATGATGAACATGAGAGACTTAAC 1268
Db 1141 TCTGGAATTTTAACTTGTGGGCGAAGACAAAGATGATGAACATGAGAGACTTAAC 1200
OY 1269 AAGCGACCTGTCTGCAAGATTCAGCACTGACCTTACGAGACAGAGATGGGAA 1328
Db 1201 AAGCGACCTGTCTGCAAGATTCAGCACTGACCTTACGAGACAGAGATGGGAA 1260
OY 1329 GGAGGCCGTGGCAATTTGATGATGCTCTTGTATGCTGACTGAGAGACTCAGCTGATCAC 1388
Db 1261 GGAGGCCGTGGCAATTTGATGATGCTCTTGTATGCTGACTGAGAGACTCAGCTGATCAC 1320
OY 1389 TTGAGACTGAGGTGTACCAAGAGCTCAAGATTTGACCTAGAGACCCACTCTTGCCA 1448
Db 1321 TTGAGAGAGAGGTGTACCAAGAGCTCAAGATTTGACCTAGAGACCCACTCTTGCCA 1380
OY 1449 GTTGTGATGATTCGAACATCTGTGAGATGCCAAATGCTTGGGCAATCAATCTGTGAT 1508
Db 1381 GTTGTGATGATTCGAACATCTGTGAGATGCCAAATGCTTGGGCAATCAATCTGTGAT 1440
OY 1509 AACATGCTGACCAATTAACCCCAAGAGCTGAATCTTCTACTAAGCGCCCAATTGGAAC 1568
Db 1441 AACATGCTGACCAATTAACCCCAAGAGCTGAATCTTCTACTAAGCGCCCAATTGGAAC 1500
OY 1569 TGGGACCAAGTGGCGGAGTGTGAGCTGAGCTGCAATGCTTGTCCACCAACGAGGGGTG 1628
Db 1501 TGGGACCAAGTGGCGGAGTGTGAGCTGAGCTGCAATGCTTGTCCACCAACGAGGGGTG 1560
OY 1629 AGCATCGAGCAGCTGACAAACGCTGCTGAGAAGCTCTAGGGGCTGTGTAACTACTCA 1688
Db 1561 AGCATCGAGCAGCTGACAAACGCTGCTGAGAAGCTCTAGGGGCTGTGTAACTACTCA 1620
OY 1689 GGGTGTCAAGTACATGTGGGCTTAATTTGTGCAAAAGAAACATGCTGCGCAAGGGCTTCTCC 1748
Db 1621 GGGTGTCAAGTACATGTGGGCTTAATTTGTGCAAAAGAAACATGCTGCGCAAGGGCTTCTCC 1680
OY 1749 TTCTGGTCTGGCTGAGCAATATCATGACCTGTGAAAGAGTATATTTGGCCCTTTGG 1808
Db 1681 TTCTGGTCTGGCTGAGCAATATCATGACCTGTGAAAGAGTATATTTGGCCCTTTGG 1740
OY 1809 AATGAAGGCTATCATGAGTGTTCATCAGCAAGAGCGGGAGCGGCCATCTAAGACA 1868
Db 1741 AATGAAGGCTATCATGAGTGTTCATCAGCAAGAGCGGGAGCGGCCATCTAAGACA 1800
OY 1869 AAGCCCCGGGCGACCTTCTCTACTGCGCTTCAAGCGAGACACCAAGAAAGAGGGGTCACT 1928
```

1801 AAGCCCCGGGACCTTCTGCTGGGTTCACTGAGACGACGAGGAAGGAGGGCTCACT 1860
1929 TTCACTTGGTGAAGAAAGACATCAGTGGCAAGACCAGATCCATCTGTAGAGCCATAC 1988
1861 TTCACTTGGTGAAGAAAGACATCAGTGGCAAGACCAGATCCATCTGTAGAACCATAT 1920
1989 ACCAAGCAGCAGCTGACACACATGTCATTTGCTGAATCATCATGGCTATTAATGATCATG 2048
1921 ACCAAGCAGCAGCTGACACACATGTCATTTGCTGAATCATCATGGCTATTAATGATCATG 1980
2049 GATGGAGACACATCTGCTGCTCCACTGCTACCTTACCTCCGACATCCCAAGAG 2108
1981 GACGCTACCAACATCTCTGTATCCCACTGCTCTACTCTTACCTGACATTTCCCAAGAG 2040
2109 GAGGCAATTTGGAAGTACTGTAGGCCCGAGAGCCAGAGCACCAGCCAGCT 2168
2041 GAGGCAATTTGGAAGTACTGTAGGCCCGAGAGCCAGAGCACCAGCT 2100
2169 AGTGTGCGCCGTACTGTAAGCAAGTTGATCTGTGACACCAAGCATGCTGACAT 2228
2101 AGTGTGCGCCGTACTGTAAGCAAGTTGATCTGTGACACCAAGCATGCTGACAT 2160
2229 ACCATTTGACCTGCGCATGTCCCGCCGACACTTATGATTCATGATGCAAGTTGGAATAC 2288
2161 ACCATTTGACCTGCGCATGTCCCGCCGACACTTATGATTCATGATGCAAGTTGGAATAC 2220
2289 GGTGAAGGTGCTGACCCCTCAGAGAGGCGAAGTTGATGCTCAGCTTACATGAT 2348
2221 GGTGAAGGTGCTGACCCCTCAGAGAGGCGAAGTTGATGCTCAGCTTACATGAT 2280
2349 CTGACCTGCGAGTGTCTACCTCCCGCATGTGAGAGAGCTGAAGCAAGCTGACAGAGAC 2408
2281 CTGACCTGCGAGTGTCTACCTCCCGCATGTGAGAGAGCTGAAGCAAGCTGACAGAGAC 2338
2409 GTGACTTGAGACACTGCCCCGTGCTCCACCCCTAAGCAGCCGAACCCCATATGCTGTA 2468
2339 -----TATGACTGAGACACTGCCATGTCACTTAGCAG-----CAACTCA 2379
2469 AACTCTTACCTTGTGTGTCAGATTTTTTTTTTAAATCTTCTACTTCTGATCTTGG 2528
2380 GATCACTGACACATCAGTGTCCAGATTTTTTCTTAACTCTCTACTGCTGCTTGG 2439
2529 GCAATCTGCGCACTTTTAAAGAGAAATGATGATGTGGGTGATTAATCTGTTATGT 2588
2440 CGAGCTGTGCTACTTTTAAATAGAACTGATGAGCGGTGGGTGATGCTCTTATGT 2499
2589 AAGAGAGAGAG-ACCTCTGAGTCTGGGATGGGCTGAGAGAGAG-----AGGGAGGCA 2641
2500 GAGGAGAGAGGACCTCTTGATCCGTGATGGCTGTGAAGAGAGAGGCTGGGGGGGAGC 2559
2642 AAGGGAGACCTCTGCTGCTGCCCCGCTGCCCTCCCTTTTTCAGAGACTCGGGGGCTTGT 2701
2560 GAAAGAGAAACACTGATGTCTGCTCGCCGTGCTCTCT-----GGTTTG 2602
2702 TTTTGAACAAGGCTCCGCTGGGCTGAGCTACCTGTTGCCCCCACTCTGTAGCTGATA 2761
2603 TTGTTGACAGAGGCTCTCTGCTG-CCATGGCAATCTGCTGCCCCCTTCTGTAGCTGATA 2661
2762 CCCCATTTCTGGAGACTCTGCTGCTGCACTTTCACCTTGTATATCCACATAGAGCT 2821
2662 CCCGAGGCTGGG-ACGCCGGCTGTGCACTTTCACCTGTTAATGTCCACATPAGAGCT 2720
2822 AGGACTAAGCCAGAGAGTCTCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2869
2721 AGGACTAAGCCATP-TAGGCTTCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2767

VERSION 188782.1 GI:3408722
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2310)
AUTHORS Kishimoto, T. and Akira, S.
TITLE Nucleic acids encoding transcription factor APRF (acute phase response factor)
JOURNAL Patent: US 5719042-A 6 17-FEB-1998;
FEATURES location/Qualifiers
source 1..2310
BASE COUNT 615 a 597 c 629 g 469 t
ORIGIN

Query Match 80.0%; Score 2294; DB 6; Length 2310;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2300; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

69 ATGGCTCAGTGGACACAGCTGTGACAGCTGGACACACAGCTACTGAACAGCTGACAG 128
1 ATGGCTCAGTGGACACAGCTGTGACAGCTGGACACACAGCTACTGAACAGCTGACAG 60
129 CTGTACAGGACAGCTTCCCATGAGCTGCGGAGTTCTTGCGACCTTGGATTGAGAGT 188
61 CTGTACAGGACAGCTTCCCATGAGCTGCGGAGTTCTTGCGACCTTGGATTGAGAGT 120
189 CAAGACTGGGCAATATGACAGCCAGCAAGAGTACATGCCAGCTGTGTTTATATCTC 248
121 CAAGACTGGGCAATATGACAGCCAGCAAGAGTACATGCCAGCTGTGTTTATATCTC 180
249 TTGGGTGAATTTGACAGCAATATAGCCGATCTGCAAGATCCCAATGCTCTATGAG 308
181 TTGGGTGAATTTGACAGCAATATAGCCGATCTGCAAGATCCCAATGCTCTATGAG 240
309 CACAACCTTGAAGATCAAGCAGTTTCTGACAGCAGGATCTTGAGAACCAATGGA 368
241 CACAACCTTGAAGATCAAGCAGTTTCTGACAGCAGGATCTTGAGAACCAATGGA 300
369 ATTTGCCGATGCTGGCCGATGCTGTGGGAAAGTCTCCGCTCCAGAGCGGACAGC 428
301 ATTTGCCGATGCTGGCCGATGCTGTGGGAAAGTCTCCGCTCCAGAGCGGACAGC 360
429 ACGGAGCCAGCAAGAGGGGCGAGGCCAACCAACAGCCGCGGTAGTGACAGAGAG 488
361 ACGGAGCCAGCAAGAGGGGCGAGGCCAACCAACAGCTGCCGTAGTGACAGAGAG 420
489 CAGCAGATTTGGAGCAGCATCTTCAAGATGTCGGGAACGAGTGCAGAGATCTAACAAG 548
421 CAGCAGATTTGGAGCAGCATCTTCAAGATGTCGGGAACGAGTGCAGAGATCTAACAAG 480
549 AAAATGAAGTGGTGGAGAACCTCCAGAGCAGACTTGAATTTCAATACAAAACCTCAAG 608
481 AAAATGAAGTGGTGGAGAACCTCCAGAGCAGACTTGAATTTCAATACAAAACCTCAAG 540
609 AGCCAAGAGACATCAGAGATCTGAATGGAACCAACAGCTGTGACACAGAGAGATG 668
541 AGCCAAGAGACATCAGAGATCTGAATGGAACCAACAGCTGTGACACAGAGAGATG 600
669 CAGCAGCTGGAACAGATGTCACAGCCCTGAGCCAGATGGGAGAGAGATTTGTGAGTAG 728
601 CAGCAGCTGGAACAGATGTCACAGCCCTGAGCCAGATGGGAGAGAGATTTGTGAGTAG 660
729 CTGGCGGGGCTTTGTCAGCAATGAGTACGTGACAGAGACACTGACTGATGAAGAGCTG 788
661 CTGGCGGGGCTTTGTCAGCAATGAGTACGTGACAGAGACTGACTGATGAAGAGCTG 720
789 GCTGACTGGAAGAGCGGCGACAGATCGCTGATCGAGAGCGCTTCCCAATCTGCTGTG 848
721 GCTGACTGGAAGAGCGGCGACAGATCGCTGATCGAGAGCGCTTCCCAATCTGCTGTG 780

QY 849 GACCGCTGGAAAACTGATTAACCTTATGACGAATCTCAACTTACACCCGCCACAA 908
DB 781 GACCGCTGGAAAACTGATTAACCTTATGACGAATCTCAACTTACACCCGCCACAA 840
QY 909 ATTAAAGAACTGGAGAGTGGACGAGAAAGTGTCTACAGAGGCGACCTTATGTCGAC 968
DB 841 ATTAAAGAACTGGAGAGTGGACGAGAAAGTGTCTACAGAGGCGACCTTATGTCGAC 900
QY 969 CACCGCCCTATCTGGAGAGAGAGATCGTGGACCTTTCAGAACTTAATGAAGATGCC 1028
DB 901 CACCGCCCTATCTGGAGAGAGATCGTGGACCTTTCAGAACTTAATGAAGATGCC 960
QY 1029 TTGCTGGTGGAGCGGACGCTTCATGCCATGCCACCGCGCCCTTATGATCAAG 1088
DB 961 TTGCTGGTGGAGCGGACGCTTCATGCCATGCCACCGCGCCCTTATGATCAAG 1020
QY 1089 ACTGGTGTCACTTTACACGAGAAAGTCAAGTTCCTGGTAAATTTCTCGAGTTAT 1148
DB 1021 ACTGGTGTCACTTTACACGAGAAAGTCAAGTTCCTGGTAAATTTCTCGAGTTAT 1080
QY 1149 CACCTTAAATTAAGTGTGATTAAGACTCTGGGGATGTGCTGCCCTCAGAGG 1208
DB 1081 CACCTTAAATTAAGTGTGATTAAGACTCTGGGGATGTGCTGCCCTCAGAGG 1140
QY 1209 TCTCGGAAATTTAATCTTGGGACAGAACACAAAGTGAATGAGAGAGTCAAC 1268
DB 1141 TCTCGGAAATTTAATCTTGGGACAGAACACAAAGTGAATGAGAGAGTCAAC 1200
QY 1269 AACGGACGCTGTCTGCAGAGATTCAAGACCTGACCTTAGGAGCAGAGATGTGGAA 1328
DB 1201 AACGGACGCTGTCTGCAGAGATTCAAGACCTGACCTTAGGAGCAGAGATGTGGAA 1260
QY 1329 GAGAGCGCGGCCAATGTGTGCTCCTTATGTGATGATGAGAGGCGACCTGATACC 1388
DB 1261 GAGAGCGCGGCCAATGTGTGCTCCTTATGTGATGATGAGAGGCGACCTGATACC 1320
QY 1389 TTGAGACTGAGTGTACACCAAGGCGCTCAAGATTGACCTAGAGCCCACTCTTGCCA 1448
DB 1321 TTGAGACTGAGTGTACACCAAGGCGCTCAAGATTGACCTAGAGCCCACTCTTGCCA 1380
QY 1449 GTTGTGTGATCTCCAACTCTGTCAATGCTTGGGCTATCAATCTGTGTAT 1508
DB 1381 GTTGTGTGATCTCCAACTCTGTCAATGCTTGGGCTATCAATCTGTGTAT 1440
QY 1509 AACATCTGACCAATAAACCAGAAAGTGAATCTTCTACTAAGCCCAATTTGAACC 1568
DB 1441 AACATCTGACCAATAAACCAGAAAGTGAATCTTCTACTAAGCCCAATTTGAACC 1500
QY 1569 TGGGACCAAGTGGCCGAGTGTCTACGTGCACTTCTGTCACCAAGCAGGCGCTG 1628
DB 1501 TGGGACCAAGTGGCCGAGTGTCTACGTGCACTTCTGTCACCAAGCAGGCGCGCTG 1560
QY 1629 AGCATGAGAGCTGACACAGCTGGCTGAGAGAGCTCTAGGGCTGGTGAATCTCA 1688
DB 1561 AGCATGAGAGCTGACACAGCTGGCTGAGAGAGCTCTAGGGCTGGTGAATCTCA 1620
QY 1689 GGGTGTCAATCAATGCGCTAAATTTCTGCAAGAAACATGCTGCAAGGCGCTTCTC 1748
DB 1621 GGGTGTCAATCAATGCGCTAAATTTCTGCAAGAAACATGCTGCAAGGCGCTTCTC 1680
QY 1749 TTGCTGGTGGCTAGACATATCATGACCTTGTGAAAAGTATATCTTGGGCGCTTGG 1808
DB 1681 TTGCTGGTGGCTAGACATATCATGACCTTGTGAAAAGTATATCTTGGGCGCTTGG 1740
QY 1809 AATGAAGGATACATGAGTTTCATGACGAGAGGCGGAGCGGCATCTAAGCACA 1868
DB 1741 AATGAAGGATACATGAGTTTCATGACGAGAGGCGGAGCGGCATCTAAGCACA 1800
QY 1869 AAGCCCCGGGACCTTCTCACTGCGCTTACGAGAGAGCAAGAGAGAGGCTCACT 1928
DB 1801 AAGCCCCGGGACCTTCTCACTGCGCTTACGAGAGAGCAAGAGAGAGGCTCACT 1860
QY 1929 TTCACTTGGGTGAAAAAGACATCACTGCGAAGACCCAGATCCAGTGTGAGACCATAC 1988

DB 1861 TTCACTTGGGTGAAAAAGACATCACTGCGAAGACCCAGATCCAGTGTGAGAGCATAC 1920
QY 1989 ACCAAGCAGAGCTGAACAACATGATTTGCTGAATCATCATGAGCTTAATGATCATG 2048
DB 1921 ACCAAGCAGAGCTGAACAACATGATTTGCTGAATCATCATGAGCTTAATGATCATG 1980
QY 2049 GATGGCAGCAACATCTCTGGTGTCTCCACTTGTCTACTCTTACCCGACATTTCCAGGAG 2108
DB 1981 GATGGCAGCAACATCTCTGGTGTCTCCACTTGTCTACTCTTACCCGACATTTCCAGGAG 2040
QY 2109 GAGCAATTTGGAAGTACTGTAGAGCCGAGAGCAGAGGACCCCGAAGCCGACAGT 2168
DB 2041 GAGCAATTTGGAAGTACTGTAGAGCCGAGAGCAGAGGACCCCGAAGCCGACAGT 2100
QY 2169 AGTGTGCCCCGATCCTGAAGACCAAGTTCATGTGTGACACCAAGACCTGACCAAT 2228
DB 2101 AGTGTGCCCCGATCCTGAAGACCAAGTTCATGTGTGACACCAAGACCTGACCAAT 2160
QY 2229 ACCATTGACCTGGCGATGTGCCCCGACCTTATGATTCATTCATGATGCAATTAAC 2288
DB 2161 ACCATTGACCTGGCGATGTGCCCCGACCTTATGATTCATTCATGATGCAATTAAC 2220
QY 2289 GGTGAAGGTGCTGAGCCCTCAGCAGAGAGGCGACGTTTGAAGTCCGCTACAGTTGACATGAT 2348
DB 2221 GGTGAAGGTGCTGAGCCCTCAGCAGAGAGGCGACGTTTGAAGTCCGCTACAGTTGACATGAT 2280
QY 2349 CTGACCTCGGAGTGTGCTTACTTCCCATG 2378
DB 2281 CTGACCTCGGAGTGTGCTTACTTCCCATG 2310

RESULT 15
AR120980
LOCUS AR120980 2787 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6159694.
ACCESSION AR120980
VERSION AR120980.1 GI:14104556
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
AUTHORS Karis,J.G.
TITLE Antisense modulation of stat3 expression
JOURNAL Patent: US 6159694-A 1 12-DEC-2000;
FEATURES
source 1. 2787
BASE COUNT 729 a 719 c 753 g 586 t
ORIGIN

Query Match 74.7%; Score 2144.2; DB 6; Length 2787;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 2305; Conservative 0; Mismatches 223; Indels 7; Gaps 2;

QY 61 GCACAGAGATGCTCACTGAGACCACTGACGACCTGACACACGCTTACCTGAAGCAGC 120
DB 213 GAAACAGAGATGGCCCAATGGAATCATGACGCTTGACACACGCTTACCTGAGCAGC 272
QY 121 TGCAACGAGCTGTACAGGACAGCTTCCCATGGAGCTGGCGCACTTCTGGACCTTGA 180
DB 273 TCCATGAGCTCTACAGGACAGCTTCCCAATGAGAGCTGGCGCACTTCTGGACCTTGA 332
QY 181 TTGAGACTCAAGACTGGGCAATGACAGCAGCAAGAGATCAATGACCTGACCTGTTTC 240
DB 333 TTGAGACTCAAGATTTGGCATATGGCGGCGAGCAAGATCAATGACCTGTTTC 392
QY 241 ATAATCTCTTGGGTAATGACAGCAATATAGCCGATTCCTGCAAGATGCAATGCTC 300
DB 393 ATAATCTCTTGGGTAATGACAGCAATATAGCCGATTCCTGCAAGATGCAATGCTC 452

```
OY 301 TCTATCAGCACACCTTTCGAGAATCAAGCACTTCTGCGAGCGAGCTATCTTGAGAC 360
    |||||
Db 453 TCTATCAGCACATCTAGAGAAATCAAGCACTTCTTCAGGCGAGTATCTTGAGAAC 512
OY 361 CAATGAAATTCGCCGATCGTGCCCGATGCTGGGAAAGTCTGCTCTCAGCA 420
    |||||
Db 513 CAATGAGATTCGCCGATGTGTGCCCGTGTGCGGAAAGATCAAGGCTTTCACAGA 572
OY 421 CGGAGCCAGCGGAGCCAGCAAGGAGGCGGAGCCAGCCAGCCAGCCAGCCAGTGA 480
    |||||
Db 573 CTGAGCCAGCTGCGGCGGAGCAAGGAGGCGGAGCCAGCCAGCCAGCCAGCTGTGA 632
OY 481 CAGAGAGGAGGAGATGTTGAGAGCAGCATCTTCAGATGTCCGGAAGCGAGGAGATC 540
    |||||
Db 633 CGAGAGAGGAGATGCTGTGAGAGCAGCATCTTCAGATGTCCGGAAGAGAGAGGAGATC 692
OY 541 TAGAAGCAAAAATGAAGTGTGTGAGAGACCTCCAGAGCAGCTTGTATTTCAACTACAAA 600
    |||||
Db 693 TAGAAGCAAAAATGAAGTGTGTGAGAGATCTCCAGAGATGCTTGTATTTCAACTATMAAA 752
OY 601 CCTCAAGAGCCAGGAGAGATGAGAGATCTGAAATGGAACCAACGATCTGTGAGCAGAG 660
    |||||
Db 753 CCTCAAGAGTCAAGAGAGATGAGAGATCTGAAATGGAACCAACGATCTGTGAGCAGAG 812
OY 661 AGAAGATGAGAGCAGTGTGAACAGATGCTCAACAGCCCTGAGCAGATGCGGAGAGCATTTG 720
    |||||
Db 813 AGAAGATGAGAGCAGTGTGAACAGATGCTCACTGCTGAGCAGATGCGGAGAGAGCATGTG 872
OY 721 TGAATGAGCTGGGCGGCTTGTGTCAAGCAATGAGTACGTCGAGAGAGACACTGACATG 780
    |||||
Db 873 TGAATGAGCTGGGCGGCTTGTGTCAAGCAATGAGTACGTCGAGAGAGACACTGACAGAG 932
OY 781 AAGAGTGGCTGACTGGAAGAGGCGGAGAGATGCGTGCATCGAGAGGCGCTCCAGACA 840
    |||||
Db 933 AGAGAGCTGGCTGACTGGAAGAGGCGGAGAGATGCGTGCATTCATTTGAGGCGGCGCCACA 992
OY 841 TGTGCTGAGACGCTGTGGAAGAACTGCATTAATCTTAGAGAGATCTCAACTTCAACACC 900
    |||||
Db 993 TGTGCTGAGTGTGCTGGAAGAACTGTGAACGTCATTAGAGAGATCTCAACTTCAACACC 1052
OY 901 GCCAACAAATTAAGAAATCGAGAGCTGAGAGAGAAAGTGTCTCAAGAGGCGAGCCTTA 960
    |||||
Db 1053 GTCAACAAATTAAGAAATCGAGAGAGTTCGACCAAAAGTTCTCTCAAAAGGAGACCCCA 1112
OY 961 TCGTGCACACCGGCGCATCTGAGAGAGAGATGCTGAGCTGTTCAGAACTTAATGA 1020
    |||||
Db 1113 TTGTACAGCACCGGCGCATCTGTGAGAGAGAGATCGTGAGCTGTTCAGAACTTAATGA 1172
OY 1021 AGAGTGCCTTCGTGTGTGAGAGCGGAGCGCGCATGCCATGCCAGCCGCGGAGCCCTTAG 1080
    |||||
Db 1173 AAGGTGCTTTGTGTGTGAGAGCGGAGCGCGCATGCCATGCCAGCCGCGGAGCCCTCG 1232
OY 1081 TCATCAGAGCTGTGTGTGCTACGATTTACAGAGAAAGTCAAGTGTCTGCTCAAAATTCCTGAGT 1140
    |||||
Db 1233 TCATCAGAGACCGGCGCTCAGTTCACACTAAAGTCAAGTGTCTGCTCAAAAGTTCCTGAGT 1292
OY 1141 TGAATTTACAGCTTAAATTAAGTGTGCAATGATTAAGAGCTCTGGGAGATGTGTGCTGCC 1200
    |||||
Db 1293 TGAATTTACAGCTTAAATTAAGTGTGCAATGATTAAGAGCTCTGGGAGAGTGTGACCTC 1352
OY 1201 TCAGAGGCTTCGGAATTTTAACATTTCTGGGAGCAGACACAAAAAGTATGAACATGAGG 1260
    |||||
Db 1353 TCAGAGGATCCCGGAAATTTTACATTTCTGGGAGCAGAAACAAAAAGTATGAACATGGAAG 1412
OY 1261 AGTCTAACACGAGCGAGCTGTCTGAGAGATTCAGAGCAGCTGAGCCTTAGGAGAGCAGAT 1320
    |||||
Db 1413 AATCTAACACGAGCGAGCTGTCTGAGAGATTCAGAGCAGCTTGAAGAGAGAGAGAT 1472
OY 1321 GTGGAGATGAGAGCGGAGCTGTGTGATGCTCTTGTGATGAGTGTGAGAGAGCTGAGC 1380
    |||||
Db 1473 GTGGAGATGAGAGCGGAGCTGTGTGATGCTCTTGTGATGAGTGTGAGAGAGCTGAGC 1532
OY 1381 TGATCACCCTTGAGAGCTGAGGTGTACCAACCAAGGCTCAAGATTGACCTAGAGACCACT 1440
    |||||

Db 1533 TGATCACCCTTGAGAGCCGAGGTGTATCACCAAGGCTCAAGATTGACCTAGAGACCACT 1592
OY 1441 CCTTGCCAGTTGTGTGATCTTCACCAATCTGTAGATGCTCCAAATGCTTGGGATCAATCC 1500
    |||||
Db 1593 CCTTGTCAAGTTGTGTGTGTCTTCACCAATCTGTAGATGCTCCAAATGCTTGGGATCAATCC 1652
OY 1501 TGTGTATTAACATGCTGTGACCAATACCCCAAGAACGTGAACCTTCTTCACTAAGCCGCAA 1560
    |||||
Db 1653 TGTGTATTAACATGCTGTGACCAATACCCCAAGAACGTGAACCTTCTTCACTAAGCCGCAA 1712
OY 1561 TTGGAACCTGGGAGCAGTGTGAGAGTGTGTGAGAGTGTGTGAGAGTGTGTGAGAGTGTG 1620
    |||||
Db 1713 TTGGAACCTGGGAGCAGTGTGAGAGTGTGTGAGAGTGTGTGAGAGTGTGTGAGAGTGTG 1772
OY 1621 GAGGCTGAGCTGTGAGAGCAGTGTGAGAGCAGTGTGTGAGAGTGTGTGAGAGTGTGTG 1680
    |||||
Db 1773 GAGGCTGAGCTGTGAGAGCAGTGTGAGAGCAGTGTGTGAGAGTGTGTGAGAGTGTGTG 1832
OY 1681 ACTACTCAGGCTGTGAGATCAGATGAGTGGCTAAATTTCTGCAAGAAACATGCTGTGCAAGG 1740
    |||||
Db 1833 ACTACTCAGGCTGTGAGATCAGATGAGTGGCTAAATTTCTGCAAGAAACATGCTGTGCAAGG 1892
OY 1741 GCTTCTCCTTGTGGGTGTGGGTGTGAGCAATATGATGAGCTTGTGAAAAAGTATATCTGG 1800
    |||||
Db 1893 GCTTCTCCTTGTGGGTGTGGGTGTGAGCAATATGATGAGCTTGTGAAAAAGTATATCTGG 1952
OY 1801 CCTTTTGAATGAAGGTATCATATGAGTGTTCATGACGCAAGAGCGGAGCGGCGCATCC 1860
    |||||
Db 1953 CCTTTTGAATGAAGGTATCATATGAGTGTTCATGACGCAAGAGCGGAGCGGCGCATCT 2012
OY 1861 TAAAGCAAAAGCCCCCGGAGCAGCTTCTGAGTGTGAGAGAGCAGAGTGTGAGAGAGG 1920
    |||||
Db 2013 TAGAGCAAAAGCCCCCGGAGCAGCTTCTGAGTGTGAGAGAGCAGAGTGTGAGAGAGG 2072
OY 1921 GGGTCACTTTCACTTGTGGGTGTGAAAAAGGAGATGAGTGTGAGAGAGCAGAGTGTGAG 1980
    |||||
Db 2073 GGGTCACTTTCACTTGTGGGTGTGAAAAAGGAGATGAGTGTGAGAGAGCAGAGTGTGAG 2132
OY 1981 AGCATTACACCAAGCAGCAGCTGTGAACAGATGATTTCTGAAATCATATGAGGCTATA 2040
    |||||
Db 2133 AACCATACACCAAGCAGCAGCTGTGAACAGATGATTTCTGAAATCATATGAGGCTATA 2192
OY 2041 AGATTCATGATGTGAGCAACATCTGTGTGTCACCTTGTGATGATGATGATGATGATG 2100
    |||||
Db 2193 AGATTCATGATGTGAGCAACATCTGTGTGTCACCTTGTGATGATGATGATGATGATG 2252
OY 2101 CCAAGAGAGAGCATTGTGAAAGTACTGTAGAGCCGAGAGAGCCAGAGAGCAGCCGAAAGCG 2160
    |||||
Db 2253 CCAAGAGAGAGCATTGTGAAAGTACTGTAGAGCCGAGAGAGCCAGAGAGCAGCCGAAAGCG 2312
OY 2161 ACCCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
    |||||
Db 2313 ACCCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2372
OY 2221 GCAGCAATACATTTGACCTGCGATGTCGCCCGGAGCTTGTGATGATGATGATGATGATG 2280
    |||||
Db 2373 GCAGCAATACATTTGACCTGCGATGTCGCCCGGAGCTTGTGATGATGATGATGATGATG 2432
OY 2281 GAAATACGCTGAGGCTGTGAGCCTCAGCAGAGAGGAGCTTGTGATGATGATGATGATG 2340
    |||||
Db 2433 GAAATACGCTGAGGCTGTGAGCCTCAGCAGAGAGGAGCTTGTGATGATGATGATGATG 2492
OY 2341 ACATGATGTGACCTGCGATGTCGCTACCTGCGGAGAGAGCTGGAACAGAGAGT 2400
    |||||
Db 2493 ACATGATGTGACCTGCGATGTCGCTACCTGCGGAGAGAGCTGGAACAGAGAGT 2552
OY 2401 GC-----AGAGACCTGACTTGTGAGACCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 2456
    |||||
Db 2553 GCAGAAAGATAGACAGTGTGAGAGCGGCTACCTGATTTGTGCGACCCCTCAGACAGCAAAACC 2612
OY 2457 CATATGCTGTGAACCTCTTACCTTGTGTGTGCTGAGATTTTGTGATTTTGTGATTTTGTG 2516
    |||||
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2002, 14:00:58 : Search time 169.97 Seconds
(without alignments)
14471.178 Million cell updates/sec

Title: US-08-212-185-11

Perfect score: 2869

Sequence: 1 GCCCGCAGCAGCCGCGCGG.....ATTATAAAAAAAAAAAAAA 2869

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

N_Genes2/1101:*

1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2869	100.0	2869	16	Mouse Stat3 (19sf6
2	2869	100.0	2869	17	Mouse Stat4 CDNA c
3	2864.2	99.8	2869	21	Mouse Stat3 nucleo
4	2418.8	84.3	2652	16	Mouse liver acute
5	2144.2	74.7	2787	15	Human placenta acu
6	2144.2	74.7	2787	21	Human low adenosin
7	2144.2	74.7	2787	21	Human Stat3 nucleo
8	2144.2	74.7	2787	22	Human signal trans
9	2132.6	72.3	3156	21	Human cancer associ
10	2074.2	74.3	2847	21	Human adenosine re
11	2012.6	70.1	2344	20	Human STAT3 allel1

12	682	23.8	2277	17	AA031278
13	680.4	23.7	2277	16	AA089338
14	672.8	23.5	2974	21	AA088863
15	672.6	23.4	2607	16	AA089337
16	672.6	23.4	2607	17	AA031277
17	672.6	23.4	2607	21	AA021053
18	672.6	23.4	2607	21	AA034931
19	672.6	23.4	3943	16	AA089336
20	672.6	23.4	3943	17	AA031275
21	672.6	23.4	4003	21	AA021052
22	672.6	23.4	4003	21	AA034930
23	672.6	23.4	4003	22	AA089227
24	672.6	23.4	7156	21	AA021058
25	672.6	23.4	7156	21	AA034936
26	619	21.6	3977	14	AA049165
27	617.4	21.5	2580	14	AA049166
28	564.4	19.7	2375	16	AA089339
29	564.4	19.7	2375	17	AA031279
30	561	19.6	1746	19	AA042537
31	550.6	19.2	2588	21	AA021236
32	550.6	19.2	2588	21	AA021237
33	550.6	19.2	2588	21	AA035114
34	550.6	19.2	2588	21	AA035115
35	550.6	19.2	2588	22	AA089230
36	550.6	19.2	2606	17	AA089324
37	550.6	19.2	2975	17	AA038325
38	550.6	19.2	8222	21	AA021239
39	550.6	19.2	8222	21	AA035117
40	448.6	15.6	519	22	AA089239
41	418.8	14.6	423	20	AA029977
42	374	13.0	423	20	AA029975
43	359.6	12.5	423	20	AA029976
44	352.4	12.3	3268	16	AA089335
45	352.4	12.3	3268	17	AA031275

ALIGNMENTS

RESULT	1
ID	AA089340
AA089340	standard; CDNA; 2869 BP.
AC	AA089340;
XX	
DT	27-SEP-1995 (first entry)
XX	
DE	Mouse Stat3 (19sf6) CDNA.
XX	
KW	Signal transducer and activator of transcription; STAT; 19sf6;
KW	Stat3; receptor recognition factor; transcription factor;
KW	cellular differentiation; derangement; dysfunction;
KW	interferon-gamma; ss.
XX	
OS	Mus sp.
XX	
FT	Key
FT	Location/Qualifiers
FT	69..2381
FT	/tag= a
XX	
PN	W09508629-A.
XX	
PD	30-MAR-1995.
XX	
PF	26-SEP-1994; 94WO-US10849.
XX	
PR	24-SEP-1993; 93US-0126588.
PR	24-SEP-1993; 93US-0126595.
PR	11-MAR-1994; 94US-0212184.
PR	11-MAR-1994; 94US-0212185.
XX	
PA	(UYRO) UNIV ROCKEFELLER.
XX	

OY 181 TTGAGAGTCAGACTGGCATATGCAGCCAGCAAGATCACAATGCCAGCTTGTTTC 240
|||||
Db 181 tlgagagtcgaagctgagcatatgcatgagccagaagatcacatgcaagcttggtgttc 240
OY 241 ATAAATCTTTGGGTGAAATTGACCACAATATAGCCGATTCCTGCAAGAGTCCAAATGCC 300
|||||
Db 241 ataattctcttggtgtaattgaccagaatalatagccgattcctgcaagatgccaatgtcc 300
OY 301 TCTATCAGCAACACTTTCGAAGATCAAGCAGTTCCTGCAGACAGATCTTGAGAAC 360
|||||
Db 301 tctatcagcaaacacttcgaaagaaatcaagcagttctgcagagcaggtatcttgagaagc 360
OY 361 CAATGGAATAATGCCCGATTCGTGCCCCGATGCTGTGGGAAGATCTCGCTCTCCAGA 420
|||||
Db 361 caatggaataatgcccgatctggtgccccgatgctgtgggaagatctcgctctccaga 420
OY 421 CGGACGCCAGCGAGCCAGCAAGGGGGCCAGGCCAACCCACAGCGCCGCTAGTGA 480
|||||
Db 421 cggacgccagcgagccagcaaggggccagggccaaacccacagccgctgagtga 480
OY 481 CAGAGAAAGCAGAGATGTTGGACAGCATCTTACAGATGTCGGAAAGCAGTCCAGATC 540
|||||
Db 481 cagaagaagcagcagatglttggagcaacatcttcaagatgtccggaagcgagtgcaagatc 540
OY 541 TAGAAGAGAATAATGAAGTGTGTGAGAACCTCCAGAGCAGCTTGAATTCACTACAAA 600
|||||
Db 541 tagaagaagaataatgaaggtgtgtgagaacctccaggaagcagcttgatttcaactacaaa 600
OY 601 CCTTCAGAGCCAGAGAGACATCAGAGATCTGAATGGAACAAACAGTCTGTAGCCAGAC 660
|||||
Db 601 ccttcagagccagagagacatcagagatctgaatggaaacaacagctcgtgaccagac 660
OY 661 AGAAGATGCAGCGAGTGGAAACAGATGCTCACACGCTCGACAGATGGGAGAGAGATTG 720
|||||
Db 661 agaagatgcagcgagctggaaacagatgctcacagcctcgacagatgggagagattg 720
OY 721 TGAATGAGCTGGCGGGGCTCTTGTTCAGCAATGAGTAGCTGCAGAAAGACACTGATG 780
|||||
Db 721 tgaatgagctggcgggctcttgttcagcaatgagtagctgcagaaagacactgattg 780
OY 781 AAGAGTGGCTGACTGGAAGAGCGGCCAGAGATGCGGTGCATCGGAGCGCTCCCAACA 840
|||||
Db 781 aagagtggctgactggaaagcgggccagagatgcggtgcatacggagcgctcccaaca 840
OY 841 TCTGCTGAGCCGCTGTGAAAACCTGATTAATTCATTAGCAGAAATCTCAACTTCAGACC 900
|||||
Db 841 tctgctgagccgctgtgaaaacctgatataatctattagcagaatctcaactcagacc 900
OY 901 GCCAACAANTTAGAAGACTGAGAGAGCTGCAGAGAAAGTGTCTTCAAGGGCGACCTTA 960
|||||
Db 901 gccacaantaatgaagactgagagagctgcagagaaagtgtctcaagggcgaccta 960
OY 961 TCGTGAGACCGCGCCCATAGCTGTGAGAGAGAGATCGTGGAGCTGTTACAGAACTTAATGA 1020
|||||
Db 961 tctgtgagacccgcccattagctgtgagagagagatcgttggagctgttcaagaacttaatga 1020
OY 1021 AAGATGCTTTCGTGTGAGAGCGGACGCTGCATGCCATGCAACCCGGACCCCTTAG 1080
|||||
Db 1021 aagatgcttctcgtgtgagagcgggacgctgcattgccatgcaacccggaccccttag 1080
OY 1081 TCAATCAAGACTGCTGTCCAGTTTACACAGCAAGTTCAGGTTTGCGTGCATTAATTCCTGAGT 1140
|||||
Db 1081 tcaatcaagactgctgtccagtttacacagcaagtctcaggttgcgtgcaaatctccgaagt 1140
OY 1141 TGAATATCAGCTTAAATTAAGTGTGATTAAGTAAAGACTCTGGGATGTTGTGTGCC 1200
|||||
Db 1141 tgaatatacagctttaaataatgaagtgtgattgataaagactctgggagtgtgtgcgccc 1200
OY 1201 TCAGAGGCTCTGGGAAATTTAACTATCTGGGCAACGAAACAAAGTGAAGACATGAGAG 1260
|||||
Db 1201 tcagaggctctgggaaatTTAACTATCTGGGCAACGAAACAAAGTGAAGACATGAGAG 1260
OY 1261 AGTCTAACACGCGACGCTGTGTGAGAGTTCAGAGCTTCAGACCTTAGCGACAGAGAT 1320
|||||

|||||
Db 1261 agtctaacaagcgacgctgtctgcagagttcaagcacctgacccttaaggagcagagat 1320
OY 1321 GTGGGAATGGAGGCGGTGCCAATTGTGATGCCCTTGTAGCTGTGAGAGGACTCCACC 1380
|||||
Db 1321 gtgggaatggagggcggtgccaatTTGTGATGCCCTTGTAGCTGTGAGAGGACTCCACC 1380
OY 1381 TGATCACTTTCGAGACTGAGCTGTACCAACCAAGGCGCTCAAGATTGACCTAGAGACCCACT 1440
|||||
Db 1381 tgaatcacttcgagacttgaggtgttacccaagcgcttcaagatttgacagagaccact 1440
OY 1441 CCTTGGCAGCTTGTGTGATCTCCACATCTGTTCAGATGCCAATTCCTTGGGCATCAATCC 1500
|||||
Db 1441 ccttggcagcttgtgtgattctccaacatctgtcagatgccaatgtcttggatcaatcc 1500
OY 1501 TGTGTATACATGCTGACCAATTAACCCCAAGACCTGAACTTCTCAGTAAGCCGCCAA 1560
|||||
Db 1501 tgtgtataacatgctgacccaataacccaagaagctgaaactcttcactaagccgcsaa 1560
OY 1561 TTGGAACTTGGACCAAGTGGCCGAGGTGCTCAGCTGGCAGTTCTGTCCACCACCAAC 1620
|||||
Db 1561 ttggaaccttggaccacaagtggtccgaaggtgtcagctggcaggtctgtccaccaccaagc 1620
OY 1621 GAAGGCTGAGCATCGAGACACTGACAAACGCTGCTGAGAACTCTAGGGGCTGTGTGA 1680
|||||
Db 1621 gaaggctgagcatcgagcagcttgacaacgctgtgtagaagctcctgaaggctgtgtga 1680
OY 1681 ACTACTAGGSGTGCAGATCATGSGCTTAATTTGTGCAAGAAACATGCTGCGACAG 1740
|||||
Db 1681 actactaagggtgtcagatcacatggtgttaaatcttgcgaagaacatgctgtgcaag 1740
OY 1741 GCTTTCCTTCTGGGCTGGCTGACAAATATCATCCAGCTGTGAAAGTATATCTTG 1800
|||||
Db 1741 gcttctccttcggtgtcgtgtgacaataatcagaccttggaaagatatacttgg 1800
OY 1801 CCTTTGGAATGAAGGATACATGCGTTTCATCAGCAAGACCGGAGCGGCCATCC 1860
|||||
Db 1801 cctttggaaatgaaggatcacatggttctacagcaaggtggtggccatcc 1860
OY 1861 TAAGCAAAAGCCCCGGGCGCACTTCTCTACTGGGCTTTCAGCGAGACACCAAGAAAGAG 1920
|||||
Db 1861 taagcaaaagccccggcgcaacttctctactgcttctcagcgagcgaaagaagag 1920
OY 1921 GGGTCACTTTCACATTTGGGTGGAAGAGACATCAGTGGCAAGACCCAGATCCAGTGTGAG 1980
|||||
Db 1921 gggctcacttctcacttgggttggaaaagaaatcagtggtgcaagaccagatccagctcgtgag 1980
OY 1981 AGCCATACCAAGCAGCAGCTGACACACATGTCATTTGCTGAATCATATGCGCTATA 2040
|||||
Db 1981 agccataccaagcagcagcttgaaacaacatgtcatttctgaaatctacatcagtgctata 2040
OY 2041 AGATCATGGATGGACCAACATCCTGTTGCTCAGTTGTATCACTTCCCGGCAATTTC 2100
|||||
Db 2041 agatcatggatggaccaaaccatcctggtgtctcactgttctacccctaccccgacatct 2100
OY 2101 CCAAGAGAGAGGCAATTTGAAAGTACTGTAGGCCCGAGACCCAGAGCAACCCGGAAGCG 2160
|||||
Db 2101 ccaagagagagagcatttggaaagtactgttaggcccgaagccagagccaccccgaaagcg 2160
OY 2161 ACCCAGGTAGTGTGCGCCCGTACTCTAAGACCAAGTTTCACTGTGTGAGACCAACGAGCT 2220
|||||
Db 2161 acccaggtagtggtgcccgtactcttgaagaccaagtctcgttggacaccaagagact 2220
OY 2221 GCAGCAATACATATTGACATGCGCATGTCCCGCGCACTTTAGATTTCATGATGAGTTTG 2280
|||||
Db 2221 gcagcaatacattgacatcctgcccgaatgtcccccgacacttgaatcatatgagcagtttg 2280
OY 2281 GAAATACGGTGAAGGTGCTGAGCCCTCAGCAGGAGGAGGAGTTTGAAGTGCATCAGTTTG 2340
|||||
Db 2281 gaataacggtgaaggtgtcgtgagccctcagcaggaaggtgagttgagtcgtccacgtttg 2340
OY 2341 ACATGATCTGACCTGTGGAGTGTGTACCTGCCCATGTGAGAGACTGGAACCAACACT 2400
|||||

Db 601 cccctaaagacaaagaaacatgcagagatctgaaatgaaacaacacgctctgacagac 660
Qy 661 AGAGATGACGACGCTGGAACAGATGCTTCACAGCCCTGGACAGATGGGAGACGATTG 720
Db 661 aagaagatgcagagcctgagacagatgctcacagccctggaccagatgacgagaagcatgtg 720
Qy 721 TCAAGTACGCTGGCGGGGCTCTTGTACGCAATGAGTACGTGAGAAAGACAGTACGTATG 780
Db 721 tgaatgaagctggcggggcctctgtcaagcaatggaaglaagtcgaagaagacatgacatgag 780
Qy 781 AAGACTGCTGCTACTGGAAGAGAGCGCCGACAGATGCTGTCATCGAGAGCGCCCTCCACAA 840
Db 781 aagagctgctgactgagaaagcgcgacagatcgctgcatcggaagcgccctcccaaca 840
Qy 841 TCTGCTGACCGCTCTGGAAAACTGATTAATCTTATTAGCAAGATTCATCTTCAACCC 900
Db 841 tctgctctgacccgctcggaaaactggatactcattgaagaatactcaactcagacc 900
Qy 901 GCCAACAAATTAAGAACTGAGAGAGCTGACGAGAAAGTGTCTCAAGAGCGGACCTTA 960
Db 901 gccacaacaatlaagaacatcgagagagctgacagagaagaatgtctctacaagggcgacccta 960
Qy 961 TCGTGCAGCAGCGGCCCATGCTGAGAGAGAGATCGTGGAGCTGTTCCAGAACTTAATGA 1020
Db 961 tctgacagacacggcccatgctggaagagagatcgtggaagctgttcagaagaactaatga 1020
Qy 1021 AAGAGCTCTTCGTTGGTGGAGCGGACCGCTGATGCCCATGACACCGGACCGGCTTAA 1080
Db 1021 aagagctctctggtggagcgagcgacgctgacatgccatgcaacccggacggcgcttga 1080
Qy 1081 TCATCAAGACTGCTGTCTCAGTTTACGACGAAGTCAAGTGTCTGTCATTAATTTCTGACT 1140
Db 1081 tcatcaagactggtgtccagtlcttaccagaagtlcaaggtgtcgtgtcaaatlctcgaagt 1140
Qy 1141 TGAATATATACGCTTAATAATTAAGTGTGATGATTAAGCTCTGGGAGTGTGCTGCC 1200
Db 1141 tgaatatacagcttaataataaagltgcatgataaagctctggggagatgtgtgcgcc 1200
Qy 1201 TCGAGAGGCTCTCGGAAATTTTACATTTCTGGGACGAAACACAAAAGTATGATGAGAG 1260
Db 1201 tcgagaggtctcggaatlaacatcttgggacgaaacacaaagatgaacatggag 1260
Qy 1261 AGCTTAACAACGCGACGCTCTGTCAGAGTTCAGACCTGACCTTAAGGAGCAGAGAT 1320
Db 1261 agcttaacaacggcagcctgtctgcagagtlcaagcacctgaccctlaaggagacagat 1320
Qy 1321 GTGGAATGAGGCGCGGCAATTTGTGATGCTTCCTTGATGCTGACTGAGAGAGCTGCAC 1380
Db 1321 gtggaaatggaagcggtgccaatltgatagtccctctgatactgtgactggaagctgac 1380
Qy 1381 TGAATCACTTCAGAGACTGAGGTGTACACCAAGGCTCAAGATTGACCTAGAGACCCT 1440
Db 1381 tgaatcacttcagagactgaggtgttacaccaagcgctcaagatltgactagagaccact 1440
Qy 1441 CATTGGCAGTTGGTGTATCTCAACATCTGTGATGCAAAATGCTTGGGCAATCAATTC 1500
Db 1441 ccttggcagttgggtgatctcaacatctgtcagaatgccaagaagcttgggcatcaatcc 1500
Qy 1501 TGTGTATTAACATGCTGACCAATTAACCCAGAAAGCTGAATCTTTCACATAAGCGCCAA 1560
Db 1501 tgtgtataaacatgctgaccaataaacccaagaagcgtgaactcttctacataagcgccaa 1560
Qy 1561 TTGGAACCTGGGACCAAGTGGCCGAGGTGCTACGCTGGCAATTTCTCTCCACACCAAG 1620
Db 1561 ttggaacctgggaccaaagtggccgaggtgctcagctgagcagtlctgtcacacaacagc 1620
Qy 1621 GAGGGCTGACATCGAGCAGCTGACAACGCTGAGTGAAGACTCTTAGGCGCTGTGTGA 1680
Db 1621 gagggctgacatcgagcagctgacaacgctgctgagaaagctctcctagggcctgtgtga 1680
Qy 1681 ACTACTGAGGGTGTGACATGACATGGCTAAATTTCTGCAAGAAACATGCTGCTGGCAGG 1740
Db 1681 actactgaggtgtcacatgacatgagatcgtcaaatctcgaagaagaacatgctgagcaag 1740

Db 1681 actactcaaggtgtcacatgacatgagatcgtcaaatctcgaagaagaacatgctgagcaag 1740
Qy 1741 GCTTCTCCTTCTGGGCTGCTGAGCAATATCATGACCTTGTGAAAAGTATATCTTG 1800
Db 1741 gcttctccttctgggctgctgagcaaatatcatcagacttgtgaaaagtatatcttgg 1800
Qy 1801 CCTTTGGAATGAAGCGTACATCATGAGTTTTCATCGAAGAGCGGAGCGGCGCATTC 1860
Db 1801 cctttggaatgaagggtacatcatatgagtttcatcgaagaagcgaggcgccatcc 1860
Qy 1861 TAAAGCAAAAGCCCCGGGACCTTCTCTACGCTTTCACGGGAGACGACCAAGAAAGAG 1920
Db 1861 taagcaaaaagccccgggacacttctcagctgtcgaagcgagcaagaagaag 1920
Qy 1921 GGGTCACTTCTCACTTGGTGGGAAAAGAGATCATGAGTGGCAAGACCCAGATCCATCTGTAG 1980
Db 1921 gggctacttctcacttgggtgggaaaagacatcgaatggcaagaccagatccagctctgag 1980
Qy 1981 AGCCATACACCAAGCAGCAGCTGACACATGCTCATTTGCTGAATCATATGCGCTATA 2040
Db 1981 agccatacaccaagcagcagctgaaacaatgcatltgtgaaatcatcatatgagctata 2040
Qy 2041 AGATCATGATGCGACCAACATCTGTGCTGCTACCTTGTCTACCTTACCCGACATTG 2100
Db 2041 agatcatgatatcgacccaacatccgtgtctcactgtgtcactgtctacacctacccgacatcc 2100
Qy 2101 CCAAGAGAGAGGCAATTTGGAAGTACTGTAGCGCCGAGAGCCAGAGACCCCGAAGCG 2160
Db 2101 ccaagagagagcatltggaagtlactgtlaggcccgaagccagagccacccgaagcgg 2160
Qy 2161 ACCCAGTACTGCTGCCCGTACTTGAAGACCAAGTTCACTGCTGTGACGCAAGCAGCT 2220
Db 2161 acccagtgatgtgctgcccgtaacttgaagacccaagttcatctgtgtgacccaagcct 2220
Qy 2221 GCAGCAATACCATTTGACCTGCCGATGTCCCGCCGACTTGTAGATTGATGATGATG 2280
Db 2221 gcagcaataacatctgacactgtcgagatgtcccccgaacttgaatltcatgtgagtlgt 2280
Qy 2281 GAAATAACGCTGAAGGCTGCTGAGCCTTCAGCAGAGAGGCGAGTTGAGCTGCTACGTTG 2340
Db 2281 gaataaacggtgaaggtgtcgaagccctcagcaagagggcagtlgtgactgtcaggtltg 2340
Qy 2341 ACATGATCTGACCTGAGCTGTGCTACCTGCCCATGTGAGGAGCTGAAGCAGAGACT 2400
Db 2341 acatgatctgacctgagctgtgctacctgcccatgtgagagctgaaacccaagagct 2400
Qy 2401 GCAGAGAGCTGACTTGAGACACCTGCCCGCTGCCACCCCTAAGCAGCGCAACCCATA 2460
Db 2401 gcagagagctgacttgagacacctgtgcccggtgtccacccctlaagcagccgaacccata 2460
Qy 2461 TCGTCTGAACCTCCTAATCTTGTGCTTCCAGATTTTCTTAAATTTCTACTTCTGCT 2520
Db 2461 tctgtctgaacctcctaatcttgtgttccagatlttlttlltaatttccactctcgtc 2520
Qy 2521 ATCTTTGGCAATCTGGGCACTTTTAAAGAGAGAAATGATGAGTGGTGTATTAAC 2580
Db 2521 atctttggcaactctggacacttlttaaaagagagaaatgagtgaggtgtgtgtataaac 2580
Qy 2581 TGTATGTAAAGAGAGAGACCTCTGAGTGTGGGAGTGGGCTGTAGAGCAAGAGGAGGC 2640
Db 2581 tgtatgttaaagagagagacactctgagtlctgtggaatggtgtgagcagaagggggagc 2640
Qy 2641 AAAGGGAACACGCTCTGCTGCGCGCGCTGCGCTCTTTTTCAGCGCTGCGGGGCTTG 2700
Db 2641 aaagggaacacactctcgtctcgtccggcctgacctcttltcaagcagctcgggggtgtg 2700
Qy 2701 TTGTTAGACAAGTGCCTCTGTTGCTCCATGCTACTGTTTGCCCATCTGTGTAGCTGAT 2760
Db 2701 ttgttagacaagtgctcctcgtgtgtccatgtctaaccgtgtgcccacatctgtgtagctgat 2760
Qy 2761 ACCCATTTCTGGGAACTCTGGGCTGTGCTGACTTTCAACTGTGCTAATTTCCATCAATGAG 2820
Db 2761 accccatctcgggaactcctgctgtgcaacttccaactgtctaatactcacatagaagc 2820

Query Match	84.3%	Score 2418.8	DB 16	Length 2652
Best Local Similarity	99.4%	Pred. No. 0		
Matches 2438	Conservative	0	Mismatches 12	Indels 2
			Gaps	1
13	CAGGCGGCGCAGTGGGCTC--AGCCGAGAGACAGTCGAGACCCCTGACTGCAGCAGAT	70		
201	cagcgccgagcagtcgggctccagcccgagagcagtcgagaccctgactgcagcagat	260		
71	GGCTGAGTGGAAACACACTCCAGCAGCTGGACACACGCTACTCGAAGCAGCTCACACGCT	130		
261	ggctcagtggaacacagctgcagcagctgcagacacagctcactcgagcagctgcacacagct	320		

QY	131	GTACACGACACGTTTCCCATGGAAGCTGGCGGACGTTCTCGGACCTTGATGATGGAGACA	190
DB	321	gtacagcaacagctctccccaatgagctcgcgagctctctgcaccttgatttgaaagca	380
QY	191	AGACTGGCATATATCACCACGCAAGAAGTCATGTCACGTTGGTGTTCATATATCTTT	250
DB	381	agactggcataatgacgacgaagaagatcaatgccaagcttggtgttcaatactct	440
QY	251	GGGTGAATTTGACCCAGCAATATAGCCGATTCTGCAAGAGTCCATGTCTCTATACGA	310
DB	441	gggtgaatttgacccaagaataatagccgattctctgcaagatccaatgtctctatcaagca	500
QY	311	CAACCTTCGAAGATTCGAAGCATTTTCGCGACACAGTATCTTGAAGAACCAATGGAAT	370
DB	501	caaccttcgaagaatccaagaagctctctgcaagcaggtatcttgaaagccaatggaat	560
QY	371	TGCCCGATCGTGGCCGATGCTGTGGGAAGTGTGCTGCTCTTCACAGCGCAGCCAC	430
DB	561	tgcccgatctggtggcgaatgctgtggaagatctgcctctcccaagcgagccac	620
QY	431	GGCAGCCGACGAAGGGGGCCAGGCCAACCCCAACAGCGCGGTAGTCAGACGAACCA	490
DB	621	ggcagccgacgaagggggccaagggccaacccaacagctgcgtatggaagaaagca	680
QY	491	GCAGATGTTGGAGCAGCATTTTCAGATGTCCCGAAGCGAGTGCAGAGCTTGAACAGAA	550
DB	681	gcagatgttggagcagcattttcagatgtcccgaaagcgagtgacagatctgaacagaa	740
QY	551	AATGAAGTGGTGGAGAACTCCAGCACGACTTTGATTCACTACAAACCTTAAGG	610
DB	741	aatgaagtgttggagaaactccagcagacttgcattcaactaaacaccttaagag	800
QY	611	CCAAATACACATGCGAGATGTCGATGTAAGAAACACCGAGTGTGACAGACACATATCA	670
DB	801	ccaaatcacatgagatctgaaatggaacaacccagctgtgacacagacgaagaatgca	860
QY	671	GCAGCTCACACGATGCTCACAGCCCTGGACCGACAGATCGGAGAACGATTGTGATGAGCT	730
DB	861	gcagctcacacgattgctcacagccctggaccagatgcygaagatctgtatgagct	920
QY	731	GGCGGGGCTCTTGTCCACCATGGAGTATCGTGGCAAGACACTGACATGGAAGACTGCG	790
DB	921	ggcggggctcttgtccacatggagtatcgtggcaagacactgacatggaaactgac	980
QY	791	TGACTGAAGAGCGCGCCAGAGATCGGTCGATCGGAGGCGCTCCCAACATCTGCTGGA	850
DB	981	tgactgaaagcgcgcagcagatctcggtcatcgagggcctcccaacatctgctgga	1040
QY	851	CCGTCGTGGAACACTGATTAATCTTATAGAGAAATCTCAACTTCAGACCCGCCAACAAAT	910
DB	1041	ccgtctgtaaacctgataacttattatgaagaatctcaacttcagaccgcgaacaat	1100
QY	911	TAAGAAATCTGAGAGAGCTGACGACGAAGTGTCTCTCAAGGGGAGACCTATGTCACAGA	970
DB	1101	taagaattctgagagagctgcaagaaanjgtctcaaaaggggagccatctgtgcagca	1160
QY	971	CCGGCCCATGCTGAGAGAGAGATCGTGAAGCTGTTTCAGAACTTATAGAGAGTGCCTT	1030
DB	1161	ccggcccatgctggaagagagatctgtgagctgtgtcagaacttaatgaagaatgctct	1220
QY	1031	CGTGGTGAAGGAGGAGCCCTGTCATGCCCATATGACCCGGAGCCGGCCCTTGTCATCAACG	1090
DB	1221	cgtagtgagagggagccctgcatagcccaatgcaaccggagcggccctatgacaaagac	1280
QY	1091	TGTGTCCAGTTTACCAAGAAAGTCAGGTTGCTGTCAAAATTTCTCGATGGAATTATCA	1150
DB	1281	tgtgtccagtttaccaagaaagtcaggtgtgtcgtcaaatcttcgtagtgtgaattacca	1340
QY	1151	GCTTAAATTTAAAGTGTGCATGTGATAAGACTGTGGGATGTTGCTGCGCCCTACAGGGTC	1210
DB	1341	gcttaaatctaaagtgtgcatgtgataagactctgtgagatgtgtgcctctaaagggttc	1400

QY 1211 TCGAATTTTAACTTCTGGGACGAAACAAAAGTATGATGACATGAGAGTCTTAACAA 1270
|||||
Db 1401 tcggaattlaacatctcgtgacgaaacacaaagtgatlaacatcggagagctlaacaa 1460
QY 1271 CGGAGGCTGTCTGACAGGTTCAGACACCTGACCTTAGGAGAGAGATGTGGGAATGG 1330
|||||
Db 1461 cggagaccctcgtcgaaggtcgaagcaccctgaaggagagagatcgtgggaatgg 1520
QY 1331 AGGCGTCCCAATTGTATGCGCTCTTGATGCTGACTGAGAGGCTGCACCTGATCACCTT 1390
|||||
Db 1521 aggcgtgccaatcgtgctcctccttgatcgtgacgtgagagcgtgaccctgacacctt 1580
QY 1391 CGAAGCTGAGGTGTACCAACCAAGGCTCAAGATTGACCTAGAGCCCACTCTTGGCAGT 1450
|||||
Db 1581 cgaagactgaagtgtaacaccaaagccccaagatgaactgaagaccacccctctgccaagt 1640
QY 1451 TGTGCTGATCTCCAAACATCTGTGAGATGCCAAATGCTTGGACATCAATCTGTGATAA 1510
|||||
Db 1641 tctgtgtatctccaacatctcgtcagatgacaaatgcttgggcatcaatccctgtgataaa 1700
QY 1511 CATGCTGACCAATTAACCCCAAGAACGTGACTTCTTCACTAAGCCGCCAATTGGACCTG 1570
|||||
Db 1701 catgtacccaataacccaagaaagtgaaactctcctaagccgcaatctgaaccctg 1760
QY 1571 GGACCAAGTGGCCGAGGTGCTCAGCTGACGTGCTCGTCACACCAAGCAGGAGGCTGAG 1630
|||||
Db 1761 ggaaccaagtggcgaagtgctcgaagctggaagctctcgtccaccaccaaagcggggctgag 1820
QY 1631 CATGAGCAGCTGACACACGCTGGCTGAGAGAGCTCCCTAAGGCTGTGTAATCTACAGG 1690
|||||
Db 1821 catcgaagacgtgacaaagcgtcgaagaaagctcctaagcgctgtgtgaactacacag 1880
QY 1691 GTGTCAGATCAATGAGGCTTAATCTGCAAGAAACATGGCTGGCAGAGGCTCTCTCTT 1750
|||||
Db 1881 gtgcagatcacaagggttaaatctcgaagaaacacatggtcgaaggctctcctcctt 1940
QY 1751 CTGGGCTGTGCTAGACAATATCATGACCTTGTAAGAAATATATCTTGACCTTTGAA 1810
|||||
Db 1941 ctggtctgtgtagaacaatcatcagaccttgtaaaaaatatactctgaccttggaa 2000
QY 1811 TGAAGGCTACATCGGTTTCATCAGCAAGGAGCGGCGCATCTTAAGACACAA 1870
|||||
Db 2001 tgaagggtacatcatgggttcatcagcaagagcggaagcgacatcccaagacaaa 2060
QY 1871 GCCCCGGGACCTTCTTCTACTGCGCTTCAGCGAGAGAGCAAGAGAGGGGTCTACTTT 1930
|||||
Db 2061 gcccccgggacacttccactcagctcagcagcagcagcaagaggaagggtctacttt 2120
QY 1931 CACTTGGGTGAAAAAGCATCATGATGGCAACCCAGATCCAGTGTAGAGCCATACAC 1990
|||||
Db 2121 cacttgggtgaaaaagcatcagtggaagaccagaatccagltctgtagagccctaacac 2180
QY 1991 CAAGCAGCAGCTGAACAACATGTATTTGCTGAATATCATGCGCTTAAGATCATGA 2050
|||||
Db 2181 caagcagcagctgaaacaacatcgtcatcttgcgaatactcatcctgaagcttaagatcatgga 2240
QY 2051 TCGCAGCAACATCTCTGCTGTCTCCACTTGTCTACTCTTACCCCGCATTTCCCAAGAGCA 2110
|||||
Db 2241 tgcgacccaacatcctcgtgtctcactgtctactccttaccctcgacatctcccaaggagga 2300
QY 2111 GGCATTTTGAAAGTACTTAGGCGGAGAGGCGAGAGCACCCTCGAAGCGAGCCGAGTAG 2170
|||||
Db 2301 ggcatttggaaagtaactcgttagagcccgagagccaaagaccccggaagccgagccaggttag 2360
QY 2171 TGTGCTCCCGTACCTGAAGACCAAGTTGATGTGTGACACCAACGACCTGACGAAATAC 2230
|||||
Db 2361 tctgtccccctgacctgaagacaaagttcatctgtgtgacacaaagacctgcaggaataac 2420
QY 2231 CATGACCTGCGCATGTGCTCCCGGACTTTGATTGATTGATGAGCTTGGAAATTAACGG 2290
|||||
Db 2421 catgacctgcgatgtctcccccgacctttagatcattgtagcagcttggaaaaaagcg 2480
QY 2291 TGAAGTGTCTGAGGCCCTCAGCAGAGAGGCGACTTGTGATGCTGACGTTTGACATGATCT 2350

Db 2481 ||||| 2540
QY 2351 GACCTCGAGATGTGCTACTCTCCCATGTAGAGAGCTGAACCAAGAGCTGACAGACGT 2410
|||||
Db 2541 gacctcgagtggtgtactcctcccatgtgagagcgtgaacccaagacgtcgaagagct 2600
QY 2411 GACTTGAGACACCTGCCCCGCTGCTCCACCCCTTAAGCAGCGCAACCCCATATC 2462
|||||
Db 2601 gacttgagacacctgcctccgtgtctccaccccttaagcagcggaaccatactc 2652

RESULT 5
AAT05616
ID AAT05616 s' .ndard; cDNA to mRNA; 2787 BP.
XX
AC AAT05616;
XX
DT 01-Apr-1996 (first entry)
XX
DE Human placenta acute phase response factor gene.
XX
KW Human; acute phase response factor; transcription factor;
KW interleukin-6; signal transmission; placenta; antibody; antisense;
KW ribozyme; antiinflammatory; antitumor; hypotensive therapy; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 221..2533
FT /*tag= a
FT /product= Acute phase response factor
FT /note= "Open reading frame (claim 7)"
XX
PN E0676469-A2.
XX
XX 11-Oct-1995.
XX
XX 29-MAR-1995; 95EP-0104670.
XX
XX 04-APR-1994; 94JP-0065825.
XX
XX (KISH/) KISHIMOTO T.
XX
XX AKira S, Kishimoto T;
XX
XX WPI: 1995-346089/45.
XX
XX P-PSDB: AAR02993.
XX
XX New acute phase response factor - for developing inhibitory agents
XX for treating diseases induced by cytokine(s) such as IL-6, e.g.
XX inflammatory diseases
XX
XX Claim 8; Page 14-15; 31pp; English.
XX
XX
CC The sequence encodes a human acute phase response factor (APRF), a
CC transcription factor related to signal transmission of interleukin-6
CC (IL-6). Human placenta cDNA is isolated using an IL-6-treated mouse
CC liver polymerase chain reaction product as a cDNA probe. APRF-
CC inhibitors, e.g. antibodies, antisense oligonucleotides or ribozymes,
CC may be used to treat diseases induced by IL-6, e.g. inflammatory
CC disease, leukemia, cancer, osteoclastia, pulmonary hypertension, etc.
XX
XX Sequence 2787 BP: 729 A; 719 C; 753 G; 586 T; 0 other;
XX

Query Match 74.7%; Score 2144.2; DB 16; Length 2787;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 2305; Conservative 0; Mismatches 223; Indels 7; Gaps 2;

QY 61 GCACGAGATGGCTCAGTGAACGAGCTGACGACGCTGACACACGCTACCTGAGCAGC 120
|||||
Db 213 gaaacagatgtcccaatgaaatcagctaacagcagcttgacacacggttacctgagcagc 272

QY	121	TGCACAGCTGTGTACAGGCACAGCTTCCCATGTGAGCTGCGGCAAGTTCTCGGCACCTTGG	180
Db	273	localcagctctaaagtagaacgtctccaatggagctctgagcagcttcttgagccctcttga	332
QY	181	TTGAGAGTCAAGACTGTGGCATATATGACAGCCAGCAAAAGTCAACATGCAGATGGTGGTTC	240
Db	333	cttgagatccaagatcttgagcatalatgcygagccagaagaatacacaatgcacatttgatgttc	392
QY	241	ATAATCTCTTGGGTGAATTGTACCACCAATATAGCCGATTTCTGCAAGAGTCCAAATGTCC	300
Db	393	ataatctctctggagagatctgaccagaglatagccgctctctgtcaagagctcgatgttc	452
QY	301	TCTATACACCAACCTTCGAAAGATTAACAGAGTTCTGCGAGAGCAGGATCTCTTGAAGAC	360
Db	453	ctctacagcaacaactctcagaagaatacaagaagcttctctcagaacagatatacttggaagc	512
QY	361	CAATGGAATTTGCCGGATCGTGGGCCGATGCTGTGGGAAGAGTCTGCGCTCCGCCGA	420
Db	513	caatggagatctgccgagatcttgagcccggtgctctgaggaagatcaacagcctctacaga	572
QY	421	CGGACGACAGCGGACGCCACAGAGGGGCCAGGCGCAACCCACACAGCGCGCCAGTGA	480
Db	573	ctgcagccacatcgcgcccgacaaggggccaagccaacaccccacaagaagcgtgagtga	632
QY	481	CAGAGAACGACAGACATCTTTGAGAGCAGCATCTTGAGATGTCCGAGAACGATGACGATC	540
Db	633	cggaggaagcagcagatgctctgagcagcagcttcagagatgtctcggaagagatgagcagtc	692
QY	541	TAGAACAGAAAAATAGCTGTGTGAGACACTCCAGAGACACTTTGATTTCAACTCAAAA	600
Db	693	tagaacagaanaatagaagtggttagagatctccaagatgactctgatttcaacttaaaa	752
QY	601	CCCTCAAGAGCCAGAGAGACATGCAGAGATCTGTAATGGAACAACAGCTGTGACCAAC	660
Db	753	ccctcaagatgacaaagagacatgcaagatcttgaatggaacaacacagatcatgtgacagtc	812
QY	661	AGAGATGACAGCAGCTGGAACAGATGCTACAGCCCTGGACACAGATGCGGAGAACATTTG	720
Db	813	agaagaatgcagcagcttggaacagatgtctcaactcgtcttgagccagatgcygagaagcatcg	872
QY	721	TGAGTGACCTGGCGGGCTCTTGTCACCAATGAGTAGACGAGAGCAACATCACTGATGTG	780
Db	873	tgagtgaagctgcygagggctcttgatcagcagtggaagtagcagcaaaactctccacgagcg	932
QY	781	AAGAGCTGGCTGACTGGAAGAGGCGGCGACAGATCCGCTGATCGGAGGCGCTCCACACA	840
Db	933	aagagctgtgactgtgaagagagcggaacaagatgtcctgcatttgagggcccgccaaca	992
QY	841	TCTGCTGTGACCGTCTGGAAAACTGGATTAATTAGCAGAAATCTCAACTTCAGACCC	900
Db	993	ctcgtctgaatctgagctagaanaactggtgaatacgtcatatlagcagaatctcacttcagacc	1052
QY	901	GCCAAACAATTTAAGAAACTGAGAGAGCTGAGCAGAGAAAGTCTCTCAAGGGCGACCCCTA	960
Db	1053	gtcaacaacaatlaagaanaactgtgagagatgtgacccaanaaagtctctacaagaagggaaccca	1112
QY	961	TCTGTCCAGCACCGGCCCATCTCTGGAGAGAGATGCTGAGAGCTTTCAGAAAATTAAATGA	1020
Db	1113	ctgtacagcacccgagcagatgtctgagagagagatctgtagagctgtgtcagaacttaatga	1172
QY	1021	AGAGTGCTTGTGTGTGAGGCGGACGCTTGATGTGCCATGACCCCGAGCCGCGCCTTAG	1080
Db	1173	aaagtgctcttggtgtgtgagcgagagccctgcatgccaatgatactctgagccgccccctcg	1232
QY	1081	TCATCAACATGCTGTCTCAGATTTACACGAAATCAAGTGTGTCGTCGTCGTCGTCGTCGTC	1140
Db	1233	tcatacaagacggggtccagcttcaactactaaagtcagtggttgctgttcaagttccctgagtc	1292
QY	1141	TGAATTTATCAGCTTTAAAAATTTAAAGTGTGATTAAGACATCTGTGGGATGTTGCTGCCC	1200
Db	1293	tgaatttcaagctttaaataatlaaagtgtgatttacaanaaactctctgagagctgttcagagctc	1352

[illegible]

Db	2433	gaataatgctgaaagctgcgcgaaccccaagcaggaaggcaglttggtcccaaccttg	2492
Qy	2341	ACATGAGTTCGACCTCGAGAGTGTGCTACCTCCGCCATGTGAGAGCTGAAACGAGAAGT	2400
Db	2493	acatgaggttgacctcgtgagctgcgtacctcccccatgtgagcagctgaaacgagact	2552
Qy	2401	GC----AGACAGCTGACTTGAGACACCTGGCCCGTGCTCCACCCCTAACAGCGGCAACC	2456
Db	2553	gcagaaagatccgactcgaagcgccctacactgcatctgcgcacccctcacagacaaacc	2612
Qy	2457	CATATGCGTCAAACTCCTAACTTGTGGTTCAGATTTTTTTTTAAATTTCCTACTTC	2516
Db	2613	cagatcacctgaaactactaactcttgtygtccaga---tttttlttaatctccacttc	2669
Qy	2517	TGCTATCTTTTGGGAATCTGGGCACTTTTAAAGAGAGAAGATGAGTGGGTGAT	2576
Db	2670	tgcatctcttgagcaatcttgcgcactttaaanaatagagaatgatgtaatgtyggtgat	2729
Qy	2577	AAACTGTATGTAAA	2591
Db	2730	ctgccttatacttaa	2744

The antisenase oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system peptide receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisenase oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisenase oligonucleotides used in the exemplification of the present invention.

Sequence	2787 BP; 729 A; 719 C; 753 G; 586 T; 0 other:
Query Match	74.7%; Score 2144.2; DB 21; Length 2787;
Best Local Similarity	90.9%; Pred. No. 0;
Matches 2305; Conservative	0; Mismatches 223; Indels 7; Gaps 2

Db 813 agaaagatgcagcagctgtaaacagatgctcactgcctgagaccagatgcggaagcaccg 872
OY 721 TGAGAGAGCTGGCGGGCTCTTGTCAGCAATGGAGTACGTACGAGAAAGCACTGACTGATG 780
Db 873 tgaatgagctgcggggcctcttgcagcagatggagtaacgctgcagaaacactcctcaagcgc 932
OY 781 AAGAGCTGGCTGACTGGAAGAGCGCGCCAGATCGCTGCATCGGAGGCCCTCCCAACA 840
Db 933 aggagctgcjctgactgtaagagcgcaacagatgtgctgcatcttggaagccgcaccaaca 992
OY 841 TCTGCTTGAGACCGTCTGGAAGAACTGGATACTTATAGCAATCTCAACTTCAGACCC 900
Db 993 tctgcctgagatcgctgtaaaactggataacgctaatgacagaaactcctaactcgaaccc 1052
OY 901 GCCAACAAATTAAGAAATCTGAGAGAGCTGCAGCAAGAAAGTGTCTACAAAGCGACCTTA 960
Db 1053 gtcaacaataaagaagaactggagaagtgacaccaaagttctcctcaaaagggagcccca 1112
OY 961 TCGTGCAGCACCGGCCCATGCTGAGAGAGATCGTGGAGCTGTTCAGAAACTTAAATGA 1020
Db 1113 ttgtaacagcacccgcccagctgctgagagagagatcgctgagacgtgttcagaaacttaatga 1172
OY 1021 AGATGCGCTTGTGTGGAGGGGAGCCCTGCTCATGCCATGCCAGCCGAGCCGCTTAG 1080
Db 1173 aaagtcctcttgctgagcgagccctgcagctcagccatgcctcagaccgcccctcg 1232
OY 1081 TCATCAAGACTGCTGTCTGAGTTTACACAGAAAGTCAGTTGCTGGTCAATTTCTGAGT 1140
Db 1233 tcatcaagaacggcgctcagctcactactactaaagtcagtgctgctgcaagttccctgagt 1292
OY 1141 TGAATTTATCAGCTTAAATTAAGTGTGATGATTAAGACTGTGGGATTTGCTGCGCC 1200
Db 1293 tgaattatcagcttaaaatlaaagtgctgcatgtaaaagactctcggaagcgttgagctc 1352
OY 1201 TCAGAGGCTCTCGGAAATTTAACTTCTGGGACAGAAACACAAAGTGTGAATGAATGAGG 1260
Db 1353 tcaagagatcccggaataatlaacatctcggcacaacaacaaagtgatgaacatggagag 1412
OY 1261 AGTGTAAACAGCGAGCGCTGTCTGTGAGATTCAAGCACCTTACCTTTGGGAGCAGAGAT 1320
Db 1413 aatcaaaaacagcagcctcctcctgcagaaatcctaactctgaccctgagggagcaagaaat 1472
OY 1321 GTGGGATGAGGCGCGTGCATTTGTATGCTCTCTTATGCTGTGATGAGAGAGCTGCACC 1380
Db 1473 gtgggaatggggcgagccaatctgtagctctccctgactgtgactgaagagcgcgcaac 1532
OY 1381 TGATCACCTTGAGAGCTGAGGTGATACGACCAAGCGCTCAAGATGACCTAGAGCCCACT 1440
Db 1533 tgaatcaaccttgagaacgagtgatatacacaaggtctcaaatctgacctlaagagaccac 1592
OY 1441 CCTTGCAGTTGTGTGATCTCCAAACATCTGTAGATGCCAAATGCTTGGCATCAATCC 1500
Db 1593 cctgtcagtgctgtagatctccaaacatctgtcagatgcacaatgacctgagcgtccatcc 1652
OY 1501 TGTGTATATACATGCTGACCAATTAACCCAAAGAGTGAACCTTCTTCACTAAGCCGCCAA 1560
Db 1653 tgtgttaacaatctgtacacaacaatcccaagaatgtgaactctctcactaagccgcaaa 1712
OY 1561 TTGGAACTGGGACCAAGTGGCGAGGTGCTCAGCTGGCACTTCTGTCGCCCAACCAAGC 1620
Db 1713 ttggaacctgggaacaaagtgcgagggcgacagctggcaatctcgcacacacaaagc 1772
OY 1621 GAGGCTGAGCATGAGCAGCTGACAACGCTGAGTGAAGAACTCTTAGGGCTGTGTGA 1680
Db 1773 gggggctgagatcagcagctgacaacgctgctggaagctcctcctgagcctggtgtga 1832
OY 1681 ACTACTGAGGCTGTCAATCATCATGGGCTAAATTTCTGCAAGAAACATGCTGCTGCAAGG 1740
Db 1833 actactaaggtgtcagatcatcatcaggtctaactctgtcaaaagaaacatggctggcaag 1892
OY 1741 GCTTCTCTTCTGGGTGCTGAGTACCAATATCATGACCTTGTGAAAAAGATATCTTGG 1800
|||||

Db 1893 gcttcctactgggtctgctgtagaacaalatcatcgacccttgtgaaagatatcttgg 1952
OY 1801 CCTTTGGAATGAAGCGGTACATCATGCGCTTTCATCAGCAAGAGCGGAGCGGCATCC 1860
Db 1953 ccccttggaagaagaggtcatcatcatatggttctcatcaagaadgagcgaggagccatct 2012
OY 1861 TAAGCACAAGCCCCGCGACCTTCTTACTGCGCTTACGCGAGAGCAACGAAGAGG 1920
Db 2013 tgaagcaactaagcccccaagcaccctctcgtgctgctcagtgtaagcagcaagagagag 2072
OY 1921 GGGTCACTTTCAGTGGGTGGGAAAGACATCAGTGGGCAAGGCCAGATCCAGTCTAG 1980
Db 2073 ggtcaacttcaacttgggtggagagacatcagcgtglaagacccaagatccaagtcgtgg 2132
OY 1981 AGCCATACCAACAGCAGCAGCTGAACATCATGCTCAATTTGTCTCAATTCATGGCTATA 2040
Db 2133 aacctatacaagaagcagcgtgaacaacaatgtaacttgcctgaaatcatcatatggctata 2192
OY 2041 AGATCATGATGCGACCAACATCTGTGTCTCCACTTGTCTACTCTTACCCGACATTC 2100
Db 2193 agatctatgtatgctaccataatcctgtgtctcactgtcatctcatctcctgacatlc 2252
OY 2101 CCAAGGAGAGGCAATTTGGAAGTCACTGAGGCCCGGAGCCGAGAGACCCCGACGCG 2160
Db 2253 ccaagagagagcatctcggaagatctgctgcgacagagacgaagagatcctgaagctg 2312
OY 2161 ACCGAGTAGTGTGCTGCCCGTACCTGAAGACCAAGTTCATGCTGTGACACCAAGACT 2220
Db 2313 acccaagtagcgctgccccatcatcctgaaagacccaattatctgtgtgacacaaagact 2372
OY 2221 GCAGCAATACCATTTGACCTGCCGATGTCGCCCGCACCTTAAATTCATTTGATGCAATTG 2280
Db 2373 gcagcaatacctatgacctgcgagctgcgcccggtcttagatcatctgctgacgttgg 2432
OY 2281 GAAATGAAGGTGAAGTCTGCTGAGCCCTCAGAGAGAGGGGACGTTTACGTCCTCAAGTTG 2340
Db 2433 gaaataatgtgaaagtgctgaaacccatcagagagagcaglttgaagttccctcaccttg 2492
OY 2341 ACATGATCTGACCTCGAGTGTGTACTCTCCCATGTGAGAGACTGAACCGAAGCT 2400
Db 2493 acatgagatgtgacctcgagtgagtgctactcccatgtgagggagctgtgaaagaaagct 2552
OY 2401 GC----AGAGAGCTGACTTGAAGACACTGCCCCGTGCTCACCCCTTAAGACCGGAACC 2456
Db 2553 gcagaaagatacagctgagcgctcactgcatctgcacacccctcacacagccaaccc 2612
OY 2457 CATATGCTGTAATAACTTACTTGTGTGTCACATTTTCTTAAATTTCTACTTC 2516
Db 2613 cagatcatctgaaactactaacttltgtgtlccaga---tlttlttaactcctacttc 2669
OY 2517 TGTATCTTTGGCAATCTGGCACTTTTAAAGAGAGAAATGATGAGTGGGTGAT 2576
Db 2670 tgcatactttagaacaatctggcaactctttaaataatagaagaatgtaagtgtgtgtgat 2729
OY 2577 AAACGTTATGTAA 2591
Db 2730 ctgcttatactaa 2744
RESULT 7
AAC93150
ID AAC93150 standard; cDNA: 2787 BP.
XX AAC93150;
AC
XX
DT 15-FEB-2001 (first entry)
XX
DE Human STAT3 nucleotide sequence SEQ ID NO:1.
XX
KW Human; mouse; STAT3; phosphorothioate; antisense oligonucleotide;
KW modulation; signal transducer and activator of transcription;
KW DNA-binding; protein; signal transduction; inhibition; apoptosis;
KW inflammatory disease;; cancer; antiinflammatory; antirheumatic;

Db 1533 tgcataccttggagaccgaggtglatcaccaggtctcaagattgacctagaagccact 1592
Qy 1441 CCTTGGCAGTTGTGTGATCTCCCAACATCTGTGATGATGCCAAATGCTTGGGCATCAATCC 1500
Db 1593 ccttgcagtggtggtgatctccaaactctgtcagatgcacaaatgctcgtgcctcctcc 1652
Qy 1501 TGTGATTAACATGCTGACCAATTAACCCCAAGAACTGAACTTCTTCACTAAGCCGCAA 1560
Db 1653 tgytgaacaacatgctgacccaacatcccaagaatgtaactcttcaactaagcgcgcaa 1712
Qy 1561 TTGGAACTTGGGACCAAGTGGCGAGTGTGTCAGCTGGCGATTCTGTGTCACCCACCAAGC 1620
Db 1713 ttggaacctgggaccagtgagccgaggtgctccagctggcctctcgtcccaaccaagc 1772
Qy 1621 GAGGCGCTGACATCGACAGCAGCTGACAAAGCTGCTGAGAGCTCTTGGGCTGTGTGA 1680
Db 1773 gggggtcgtgagcatcgagcagctgacaaagctgctgagaaagctcctcctggtcgtgtga 1832
Qy 1681 ACTACTAGGGGTGTGATGATCATGAGGCTAAATTCTGCAAGAAACATGGCTGGCAAGC 1740
Db 1833 actactaaggtgtcagatcacatggtctaactctgcaagaacacatgctgcaag 1892
Qy 1741 GCTTTCCTTCTGTGGTGTGGCTGACATATATCATCGACCTTTGTGAAAAGATATCTTGG 1800
Db 1893 gcttctcctcagtggtctgagcaaatatcatcgaccttggaataaagatatactctgg 1952
Qy 1801 CCCTTTGGAATGAAGGCTATCATAGGGTTTCATCAGCAAGAGAGCGGAGCGGCATCC 1860
Db 1953 ccccttggaatgaaggtatatacatactggttctcagcaagagcggtgcggtcactc 2012
Qy 1861 TAAGCAAAAGCCCCCGGACCTTCTACTGGCTTCAGCGAGAGCAGCAAGAAAGAGAG 1920
Db 2013 tggacctaagcccccagagccctctcgtcgtcctcagtgaaagcagcaaaagagag 2072
Qy 1921 GGGTCACTTTGCTGGTGGTGAAGAGACATAGTGGCTTCAGCGAGAGCAGCAAGAAAGAGAG 1980
Db 2073 ggtcacttctcacttggtggtgagaagacatcagcgttaagaccagatccagtcgtgg 2132
Qy 1981 AGCCATACCAAGCAGCAGCTGAACAACATGTCATTTGCTGAATCATGAGGCTATA 2040
Db 2133 aaccatacaaaagcagcagctgacaacatgctcatctgtaaatcatcatctgggtcata 2192
Qy 2041 AGATCATGATGGAGCAACATCTGTGTCTCCACTTGTCTACTCTTACCCGACATTC 2100
Db 2193 agatcctgagtgctaccataatccctgtgtctcactgtctatctatctatcagacatc 2252
Qy 2101 CCAAGAGAGAGGAGCTTTGAAAGTACTAGGCCGAGAGCCGAGAGAGAGAGAGAGAGAGAG 2160
Db 2253 ccaagagagagagcattcggagagatattgttcgccaagagagccagagcattcgtgaagctg 2312
Qy 2161 ACCCAGGATGCTGCTGCCCGCTACCTGAAGACCAAGTTTCATCTGTGTGACACCAAGACT 2220
Db 2313 acccaggtagcgtcccccctacatctgaaagcagaattatctggtgaaccacaagcct 2372
Qy 2221 GCAGCAATCCATTAAGCTGCGATGCTGCCGATGCCCGCACTTAAATTTATGATGAGTTTG 2280
Db 2373 ggcagcaatacattgacacctgagcagatgcctcccgcgctttagatctatctatgtagagttg 2432
Qy 2281 GAAATTAAGGTAAGGTGCTGAGCCCTCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
Db 2433 gaataatggtgaggtgctgaaacctcagcagagaggtgaggttgcctcctcaacttg 2492
Qy 2341 ACATGATCTGACCTGGAGAGTGTGTAACCTCCCATGTGAGAGAGCTGAACCAAGAGCT 2400
Db 2493 acctgagttgacctgaggtgctgtaacctcccccattgtgagagagcgtgagaaagagct 2552
Qy 2401 GC----AGAGACGTGACTTGAACACCTGCCCGTGTGCTCAACCCCTTAAGAGCGGACCC 2456
Db 2553 ggcgaagaagatacagcagaggtgcctcactgcatcttcgcccacctcaacacagccaaacc 2612
Qy 2457 CATATGCTGTAAGACTCTCACTTTGTGTTGCAGATTTTATTTTAAATTTCTACTTTC 2516
Db 2613 cagatcatctgaaactaactactgtgtgtccaga---tttttttaactctcctacttc 2669

Qy 2517 TGTATCTTTGGGCAATCTGGGACCTTTTAAAGAGAAATGATGATGTGGTGCAT 2576
Db 2670 tgcataccttggagcaatctgggacactttaaataagaaatgagatgtagatgtgggtgat 2729
Qy 2577 AACTGTTATGTATA 2591
Db 2730 ctgcttattctctaa 2744

RESULT 8
AAAB9229 standard; cDNA: 2787 BP.
ID AAA89229:
AC AAA89229:
XX
DT 28-MAR-2001 (first entry)
XX
DE Human signal transducer and activator of transcription STAT-3 DNA.
XX
KW STAT-3; signal transducer and activator of transcription 3;
KM human; crystal; drug screening; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 221..2533
FT /*tag= a
XX
PN US6160092-A.
XX
PD 12-DEC-2001
XX
PF 29-MAY-1998: 98US-0087465.
XX
PR 29-MAY-1998: 98US-0087465.
XX
PA (UYRQ) UNIV ROCKEFELLER.
PI Chen X, Darnell JE, Kuriyan J, Vinkemeier U, Zhao Y, Jeruzalmi D;
XX
DR WP: 2001-101568/11.
DR P-PSDB: AAB19964.
XX
PT Novel crystal useful in drug screening assays, comprises portion of
XX signal transducer, activator of transcription and duplex DNA
XX
PS Disclosure: Column 63-68; 206pp; English.
XX
CC The present sequence is that of DNA encoding human signal
CC transducer and activator of transcription 3 (STAT-3, see AAB19964).
CC The invention provides a crystal of a core portion of a STAT
CC protein in dimer form with an 18-mer duplex DNA (see AA89233) that
CC contains a binding site for the STAT dimer. The core portion
CC comprises an N-terminal alpha-helical domain, a DNA binding domain
CC which contains an immunoglobulin-like fold, a C-terminal SH2 domain
CC and a domain that links the DNA binding and SH2 domains. The
CC crystal is of sufficient quality to perform X-ray crystallography
CC studies. Methods of preparing the crystals are included in the
CC invention. Knowledge of the STAT protein's 3-dimensional structure
CC will aid in structure-based drug design. The crystal can be used
CC in drug screening assays to identify agonist and antagonist
CC compounds. Antagonists can be used to treat inflammation, allergy,
CC asthma and leukaemia, and agonists to treat anaemia, neutropenia,
CC thrombocytopenia, cancer, obesity, viral diseases, growth
CC retardation, and other conditions characterized by insufficient
CC STAT activity.
XX
SQ Sequence 2787 BP; 729 A; 719 C; 753 G; 586 T; 0 other;

Query Match 74.7%; Score 2144.2; DB 22; Length 2787;
Best Local Similarity 90.98; Pred. No. 0;

Matches 2305: Conservative 0: Mismatches 223: Indels 7: Gaps 2:									
Oy	61	GCAGAGATGGCTGCTAGTGAACACGCTGCACAGCTGACACGCTACTGTAACAGC	120						
Db	213	gaaacagagatgcccacaaagaaatcagctacagcagcttgacacacggtaccctggagcagc	272						
Oy	121	TGCACCAAGCTGTACAGCAGCAGCTTCCCATAGAGCTGGCGCAGTTTCTTGACACTTGG	180						
Db	273	tcacatcagctctacagctgacagcttcccaatgagctgcgagcttctcggcccttgg	332						
Oy	181	TTCAGAGTCAACAGCTGGGATATGACGACGAAAGATGCATGCGACGTTGGTGTTC	240						
Db	333	ctgagagatcaaatgctggacatactggccgaagaaatcaatcacttgggtttc	392						
Oy	241	ATAATCTCTTGGGTGAATTTGACCAAGTAATAGCCGATTCCTGCAAGAGTCCATGTCC	300						
Db	393	ataatctcttggagagatctgaccagcagatagccgcttccgcaagctcgaagctgaatgctc	452						
Oy	301	TCTATCAGACAACTTCCAAAGATACAGCACTTCTGCAGACGAGTATCTTGAGAACG	360						
Db	453	tcatacagacaactctacgaagaaatcaagcagttctctcagagcaggtatctgagaagc	512						
Oy	361	CAATGAAATTTGCCGATCGTGGCCGATGCTGTGGAGAGTGTGCGCTCTCCAGA	420						
Db	513	caatgagatctgccggtatctgtgcccgtgtcctgtgggaagaaatcaagccttctacaga	572						
Oy	421	CGGACGCCAGCGACGCCAGCAAGGGGGCCAGCCACACCCACAGCCCGCTAGTGA	480						
Db	573	ctgacgcaactgcgcccagcaaggggccaagccaaccccccacagcagcgctggltga	632						
Oy	481	CAGAGAAAGCAGATGTTTGAGCAGCATCTTACAGATGTCCGAGCGGAGCGGACAGATC	540						
Db	633	cgaagaaagcagcagatgctggagacaacctcagagatgctcggaaagagatgcaagatc	600						
Oy	541	TAGAAGAAAGAAAGAGTGTGAGAACCTCCAGACAGCACTTGTGATTTCACTACAAA	600						
Db	693	tgaagacaaaatgaaagtggtagagaatctccaagatgacttgatcttaactataaaa	752						
Oy	601	CCCTCAAGAGCCAGAGACATCGAGATCTGTAATGGAACAAACAGTCTGTGACAGAC	660						
Db	753	ccctcaagatccaagagacatgcaagatctgaatgaaacaaacagctcagtgccaagc	812						
Oy	661	AGAGAGTGCAGCGCTGGAGACAGATGTCACAGCCCTGGACCGCATCGGAGAAAGCTTG	720						
Db	813	agaagatgagcagctggaaacagatctcacgtgcgtggaccagatggcggagaagcctcg	872						
Oy	721	TGAGTGAAGTGGGGGGGCTTGTTCAGCAATGGATGAGTACGTGCAAGAGACACTGATATG	780						
Db	873	tgagtgagctggcgggcttcttgcaagcgaatgagtaatgctgcagaaactctcaagcga	932						
Oy	781	AAGAGCTGGCTGACTGGAAGAGCGGCCAGAGATCGGTGCATCGAGGCCCTCCACA	840						
Db	933	aggagctggtgactggaaagagcgcaacagatgctgcatltagagggccgcacaca	992						
Oy	841	TTCGGCTGGACCGTCTGGAAAGCTGATACTTCAATGAGAGATGTCACACTTCAGACC	900						
Db	993	ctcgcccaagctcgagctagaaaactgtaacgltcaatagaagaatcccaatcaagccc	1052						
Oy	901	GCCAAACAAATTAAGAAACTGAGAGAGCTGCAGCAAGAAAGTGTCTTACAGAGCGACCTTA	960						
Db	1053	gtcaacaatataagaaacttgaggagtgtgcacaaaagtctcttacaagaagggaaccca	1112						
Oy	961	TGCTGCAGCACCGGCCCATCTGTGAGAGAGATCGTGGAGCTGTTGAGAACTTAATGA	1020						
Db	1113	ctgtacagcacccgcccagatgctggaagagagatcgtggagctgtltaagaaactaatga	1172						
Oy	1021	AGAGTGCCTTGTGGTGGAGCGGACCGCTGCATGCCATGCAACCGGACGCGCCTTAG	1080						
Db	1173	aaagtgcccttgtagggagcgagccctgcatgcccacatgcatccgagccgcccctcg	1232						
Oy	1081	TCATCAAGACTGTGTGCTACGTTACAGCAAGAGTCAAGTGTGCTGGTCAATTTCTGAGT	1140						
Db	1233	tcatacagaccggtctcagcttcaactactaaagtcaggtctgctggtcaagttcccttgat	1292						

Oy	1141	TGAATTATCAGCTTAAATTTAAAGTGTGCATTTGATATAAGACTGTGGGATGTTGCTGCC	1200						
Db	1293	tgaattatcagcttaaatataaagtgtgcatlgaacaagaactctgggaagctgtgcagctc	1352						
Oy	1201	TCAGAGGCTGTGGCAATTTAACATTCCTGGCAGCAGAACACAAAAGTATGATGAGTGA	1260						
Db	1353	tcagagagatccggaaatcttaacatctcggccacaacaacaagaatgataatgaatgaag	1412						
Oy	1261	AGTTCACCAACGCGCAGCTGTCTGCAGAGTTCAAGCAGCTGACCTTAGGAGCAGAGAT	1320						
Db	1413	aatccaacaacgacgctctctcgcaagatctcaaacacttgacccttgaggagcagaagat	1472						
Oy	1321	GTTGGAATGAGGCGCGTCCCAATTTGATAGCTTCCTTGATGCTGACCTGAGAGCTGACC	1380						
Db	1473	gtggaaatggggccgagccaatgtgatagtctccctgattgtgaatcgaagagctgcacc	1532						
Oy	1381	TGATACACTTCGAGACAGAGGTGATACCAAGGCTTCAAGATTGACCTAGAGACCCACT	1440						
Db	1533	tgatacactcttgagaccggaagtgtatcaccaaggtctccaagatgtgacctagagaccact	1592						
Oy	1441	CTTTGCCAGTTGTGTGATCTCCAAACATCTGTGCAGATGCCAAATGCTTGGCATCAATC	1500						
Db	1593	ccctgtcagcttggtgtgatactccaacactcgtcagatgccaatgacctgggctgctcacc	1562						
Oy	1501	TGTGCTATTAACATTCCTGACCAATTAACCCCAAGACGTAACCTTGTACTAAGCGGCA	1560						
Db	1653	tggtgatacaatctgacccaacaaatcccaagaatgtgaactctctcactcaagccgcaaa	1712						
Oy	1561	TTGGAACCTGGAGACCAAGTGGCGCAGGTGCTACAGCTGGAGTTTCTCTCCACCAAGC	1620						
Db	1713	ctggaacctgggaccagaatggccgaggtgctcagctggcagttctcgtccaccaaccaagc	1772						
Oy	1621	GAGGCTGAGCAGTACGACACCTGCAACCGTGGCTGAAGAACTCTTAGGGCTGTGTTGA	1680						
Db	1773	gggggctcgaatcagcagcagctgcaacagctgctgtagaagcttcccttaggscctgtgtga	1832						
Oy	1681	ACTACTCAAGGTGTACATACATGAGGCTAAATTCCTCAAGAAAGAAACATGGCTGCAAG	1740						
Db	1833	actactcaaggtctcagatcaatcaatcagctgcaacactctcgaagaacaatgctggcaag	1892						
Oy	1741	GCTTCTCTTCTGGGTGTGGCTAGACAAATATATCATGCACTTGTGAAAAGCTATATCTTG	1800						
Db	1893	gcttctcctactgggtctgctgtaagaataatcatcatcagacacttgtaaaaagataatcttg	1952						
Oy	1801	CCCTTGTGAATGAAGGATACATCATGGGTTTCATACAGCAAGCAGCGGAGCCATCC	1860						
Db	1953	cccttgaaatgaagggtatcatcatctggtttcatcagaagacggaagcgaggccatctc	2012						
Oy	1861	TAAACCAAAAGCCCCCGGACACTTCTACTGCGCTTACGAGGAGAGCAAGCAAAAGAG	1920						
Db	2013	tgagcaactaaagccccagggcaactctctgctgagcttcaatgaaagcagaagaagagag	2072						
Oy	1921	GGGTCACTTCTACTTGGGTGGAAGAGCATCACTGCGCAAGACCCAGATCCAGTCTTAG	1980						
Db	2073	ggctcacttctcaacttgggtggagaaagacatcagcggtlaagaccagatccaagctcgltg	2132						
Oy	1981	AGCCATACACCAAGCAGCAGCTGTAACAACATGTCATTTGCTGAAATCATCATGGGCTATA	2040						
Db	2133	aacctacacaagaacagcagctgaaacaacatgtcatctgtcgaatcatcatatggtgcata	2192						
Oy	2041	AGATCATGATGAGGAGCAACATCTCTGGTGTCTGCATTTGTACTCTACCCGCACTTC	2100						
Db	2193	agatcatggaatgactaacaatatctgtgtctccaactgtctatctctatctcgtgaatc	2252						
Oy	2101	CCAAGGAGAGGCAATTTGGAATGACTGTAGGCCGAGAGCCAGAGACACCCCAAGCCG	2160						
Db	2253	ccaagagagaggaatctgggaagatattgtcgccaagagagccaagatccctgaaagctg	2312						
Oy	2161	ACCAGAGTAGTGTGCCCGCTACTCTGAAGCAACAAAGTTATCTGTGTGACACCAAGCACT	2220						
Db	2313	accaggtlaagctgcccacatcactgaagaccagttatctgtgtgataccaccaagacct	2372						

QY 2221 GCACCAATACCATGACCTGCCGATGTCGCCCGCACTTAGATTGATTGATGAGTTTG 2280
|||||
Db 2273 gcagcaataccattgacctgcgagtgcccccgcgtcttagatcaattgagcagttg 2432
QY 2281 GAATAACGCTGAAGGCTGCTGAGCCCTCAGCAGAGAGGCACTTGTGACGCTCAGCTTGG 2340
|||||
Db 2433 gaataatggaaggtgctgaaccctcagcaaggaaggaagcttgaagctccctcagcttg 2492
QY 2431 ACATGATCTGACCTCGAGCTGCTACTCTCCCATGTGAGAGAGCTGAACAGAGAGCT 2400
|||||
Db 2493 acatgaggttgacctgagtgagctaccctcccatgtagagagctgaggaagcgaagct 2552
QY 2401 GC---AGAGCTGACTTGAGACACTGCCCGCTGCTCCACCCCTAAGCAGCGCAACC 2436
|||
Db 2553 gcagaaagatcagactgagcgctaccctgactctgctccaccctcacacagccaaaccc 2612
QY 2457 CATATCGCTGAACTCTACTTGTGTTCCAGATTGTTTATTTTATTTCTACTTTC 2516
|||||
Db 2613 cagatcaatcgaactactaactcttggtgccaga---tttttttaactctcactc 2669
QY 2517 TGTATCTTTGGCAATCTGGCAGCTTTTAAAGAGAGAAATGAGTGTGGGTGAT 2576
|||||
Db 2670 tgcatactctgagcaatctggcactcttaaaatagagaatgagatgagtggtgag 2729
QY 2577 AACTGTATGTATA 2591
|||
Db 2730 ctgcttactcttaaa 2744

RESULT 9
AAFI8318
ID AAFI8318 standard; DNA; 3156 BP.
XX
AC AAFI8318:
XX
DT 14-MAR-2001 (first entry)
XX
DE Lung cancer associated polynucleotide sequence SEQ ID 337.
XX
KW Human: lung cancer associated protein; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active; vulnerable;
KW gastrointestinal; nephrotoxic; anti-infective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease; ds.
XX
OS Homo sapiens.
XX
PN W020005180-A2.
XX
PD 21-SEP-2000.
XX
PE 08-MAR-2000; 2000MO-US05918.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PA (ROSE/) ROSEN C A.
XX
PI Ruben SM:
XX
DR WPI: 2000-587514/55.
XX
DR P-PSDB: AAB58442.
XX
PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -
XX
PS Claim 1; Page 794-795; 1425pp; English.
XX
CC Polynucleotide sequences AAFI7982 - AAFI8424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;

CC immunomodulatory; muscular active general; vulnerable; gastrointestinal
CC general; nephrotoxic; anti-infective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer,
CC chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAFI8425 - AAFI8433 and
CC peptide AAB58549 are used in the course of the invention for the
CC identification and characterization of the polynucleotide and protein
CC sequences.
XX
SQ Sequence 3156 BP; 832 A; 792 C; 826 G; 705 T; 1 other:
Query Match 74.3%; Score 2132.6; DB 21; Length 3156;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 2464; Conservative 0; Mismatches 334; Indels 51; Gaps 8;
QY 68 GATGGCTCAGTGAACCACTGACAGACTGGACACAGCTACTGAGCAGCTGACCA 127
|||||
Db 194 gatgcttaatgataatcagctacagagcttgacacagtaactgagcagctccatca 253
QY 128 GCTGTACAGCCACAGCTTCCCATGAGAGCTGGGACGCTCTGAGCTTGTGAGAG 187
|||||
Db 254 gctctacagtgacagctcccaatgagctgagcagcttctgccccttgagtgag 313
QY 188 TCAAGACTGGGCAATGAGCAGCAAGAGTGCATGCGACGTTGGTGTATTAATCT 247
|||||
Db 314 tcaaatgagcatatgagcagcagcaagaatcacatgcccacttggttctaatc 373
QY 248 CTTGCTCAATGACGACATATAGCCGATTCCTGACAGAGTGCATGCTCTATCA 307
|||||
Db 374 cctgctcaatgacagcagctatagcgccttcttcaagagctgagatgctctatca 433
QY 308 GCACACCTTGGAGAAATCAAGCAAGTTCTGACAGCAGTATCTTGAAGCCCAATGA 367
|||||
Db 434 gacaaatcaagaagaatcaagaagcttctcagaagcagatcttgagaagccaatgga 493
QY 368 AATTGCCGGATCTGTGGCCCATGCTGTGGAAAGTCTGCTCTCCACGCGAGC 427
|||||
Db 494 gatgcccagattgtgcccgtgcttgggaagaatcacagcctctacagactcagc 553
QY 428 CACGCGACCCGACGAAGGGGCGCCACGACCAACCCGCGATGACAGAGAA 487
|||||
Db 554 cactgcccagcaagaaggcgccagccacccacagcagcggctgagcagga 613
QY 488 GCAGCAGATGTTGGAGCAGCATCTTACAGATGTCGGAAGCGAGTGCAGATAGACA 547
|||||
Db 614 gcagcagatgcttgagcagcagcactcagatgctcggagaagagtgagcagatcaga 673
QY 548 GAATAATGAAGTGTGGAGAACCTTCAGAGAGCTTTGATTCAACTCAAAACCTCAA 607
|||||
Db 674 gaaatgaagtgtagaagaatctcagagagacttgattcaactataaaacccctcaa 733
QY 608 GAGCCAGAGCATGACGAGATCTGAATGAACACACAGCTCTGACACAGAGAGAT 667
|||||
Db 734 gaggcaagaagacatgcaagatctgaatggaacaacacagcagatgacagcagcagaat 793
QY 668 GCAGCAGCTTGAACGATGCTTCACAGCCCTGACCCAGATGCGGAGAAAGCATTTGAGTGA 727
|||||
Db 794 gcaagcagctggaacagatgctcactgctgctgacacagatgcggaagcagcagtgag 853
QY 728 GCTGGCGGAGCTCTTGTACGATGATGATGATGATGATGATGATGATGATGATGAT 787
|||||
Db 854 gctggcggagcttctgtagcagatgagtagcagcagaaactctcagcagcagagcct 913
QY 788 GCGTCACTGGAAGAGCGCGCAGAGATGCGTGCATGAGAGCCCTCCACATCTCCCT 847
|||||
Db 914 gctgactcaggaagggcgcaacagatgctgacttgaggaagccgcgccaacatctgct 973

QY	848	GGACCGCTCTGGAACAACTGGATTAACCTTCATTAGCAGAAATTCACACTTCAGACCCCGCCANCA	907
Db	974	agactcgcgtctgaagaaactctgataaactcgtatctatgagaatctcaactctcagaccgcgtcaaca	1033
QY	908	AATTAAAGAACTGGAGAGACTGCAGCAAGAAAGTGCTCTCAAGGGCGACCCATTCGTCA	967
Db	1034	aattaaagaactcggagagctctgacgaaaaaagcttctccacaaggagggcccatctgtaca	1093
QY	968	GCACCGCCCATGCTGGAGAGAGAGATCGTGAGCTGTTCCAGAAACTTATGACAGAGTC	1027
Db	1094	gcaccgcgcgcagctcgtgagagagaaatcgtlgagcgcgttcttaagaacttaatgaaagctgc	1153
QY	1028	CTTCGTGGTGGAGCGGACCGCTTCGATGCCCATAGCAAGCCGGACCGCCCTTAGTCATAA	1087
Db	1154	ctctgtgtgtgagagcgagacccctcgacagcccaatgcaatccctgcagcggccctctgtacaa	1213
QY	1088	GACTGTGTCCAGTTTACCCAGAAAGTCAGGTGCTGCTGCAATTTCCGATTTGAATTA	1147
Db	1214	gaccgcgcgtccacgtctactactaagctacagtgctgtgtgtcgtcaaatctccctgagttgaata	1273
QY	1148	TCACGTTTAAATTTAAAGTGTGCAFTTGATAAGACTCTGGGGATGTTGCTGCCCTTACAGG	1207
Db	1274	tcagctcttaaaatlaaagctgtgcatctgacaagaactcttgggagcgtgtgcagctctcaagag	1333
QY	1208	GTCCTCGAAATTTAACATTTGTGGGACGAGAACACAAAGGATGAACATGAGAGACTTCA	1267
Db	1334	atcccgagaacttaactacatctcttgggcccacaacacaaaagctgatacagatacgtgaatccaa	1393
QY	1268	CACGCGCAGCCTGTCTGCAGAGATTCAAGCACTGCACCTTTAGGGAGCAGAGATGTGGAA	1327
Db	1394	caacgcgcagcctctctctgcagaatctcaaacacttgacccttgagagagcagaagatctgaggaa	1453
QY	1328	TGCGAGCGCTCCCAATTTGTATGATGCCTTCCTTGATCGTGATGGAGAGCTGCACCTGATAC	1387
Db	1454	tgggggcccgaagccaattgtgatagtctctccctgtatctgtacatcggagagcctgcacgtgatac	1513
QY	1388	CTTCGAGACTGAGGTGTACCCAGCCAGAGCCCTCAAGATTGACCTAGAGACCACCTCTTGCC	1447
Db	1514	cttgaagaccgaagtgtatcatccaagccctcaagatctgacccaagaccccaactctctgc	1573
QY	1448	AGTTGTGGTATCTCCCAAACTCTGTAGATGCCAAATGCTTGGGCATCAATCTTGCTGTA	1507
Db	1574	agctgtgtgtctcccaactcctgtcaagatgcacaatgcctgggagctcatctctgtgtgta	1633
QY	1508	TACATGCTGCACAAATAACCCCAAGAACGTGAATCTTCTCAGCTAAGCGCCGCAATTGGAAC	1567
Db	1634	caactcgtctgacacaacaatcccaagaatgtaactcttctaccagccccccaattgtgac	1693
QY	1568	CTGGGACCAATGGCCGAGAGTGCCTCAGCTGTGCAAGTTCTGCTCCACCACCCAGCGAGGCT	1627
Db	1694	ctgtgataaagctgtgcgaagctcctgtcagctgtgcagctctctccaccacaacgaacgagact	1733
QY	1628	GAGCATGAGAGAGCTGCACACGCTGGCTGAGAGCTCTAGGGCTCGTGGTGAATTAATCT	1687
Db	1754	gagatctcgagagcgtgtactactactgtgcagagaacactcttggagactcgtgtgtgataatctc	1813
QY	1668	AGGGTGTACATACATGAGGCTTAAATTTCTGCAGAAAGAAACATAGCTGCGACAGGCTTCTC	1747
Db	1814	aggggtgtcaagatcacatcgtggtctaattcttgcagaagaacatcgtctgcgaagggctcttc	1873
QY	1748	CTTCGGGTGTGGCTAGACAAATATCATATCCAGACCTTGGAAGAAATATATCTTGCCCTTTG	1807
Db	1874	cttcgcgtgtcgtcgtgacataatcatctgacctcttgcgaagaagatcatctcgtgcctcttg	1933
QY	1808	GAATGAAGAGTACATCTGGGTTTCTACAGCAAGAGAGCGGAGCGGCAATCTTAAGAC	1867
Db	1934	gaacgaagggtacatactatgggcttatacgtlaagagcgggagcggtcatcttgagac	1993
QY	1868	AAACCCCGCCGACACTTCTCTACTGCGCTTCAAGCAGACAGACCAAGAGAGGCGCTTAC	1927
Db	1994	taagctctccaggaacctctctgtctaagatctcagtgtaaaagcgaagaagagagcgtctac	2053

QY	1928	TTTCACCTGGGGTGGAAAAAGCAGCATCACTAGTGGCAAGACCACCGAGATTCACACTGTGTAGACCCCTA	1987
Db	2054	tttcacttgggttggagaagaagacatcaacggtaagaccacagatccagttccgttggaaacca	2113
QY	1988	CACCAGCAGCAGCGCTGAACCAACATGTCATTTGGCTGAATATCATCATGCGCTATTAAGATCAT	2047
Db	2114	cacaaagcagcagcttgaacaaacatgtaacttltgttgtaaatcalcatatgggtataaagatcat	2173
QY	2048	GGATCGCAGCAACATCCTGGTGTCTCCACTTGTCTACCTCTACCCCGACATTTCCCAAGCA	2107
Db	2174	ggatgcctacaaatccctgtgtctccactgtctactctatccctgaatctcccaaggaa	2233
QY	2108	GGAGGCATTTTGAAAGTACTGTAGCGCCCGAGACCCGAGGACGCCGCCGAAGCCGACCCAGG	2167
Db	2234	ggaaggacttcggaaagtaatctgtcggccagagagccagagacatccgaaactgtaccacag	2293
QY	2168	TAGTACTCTCCCCCTACCTGGAAGACCAAGTTTCACTCTGTGTAGACACCAACCACTTCACGCA	2227
Db	2294	taggctctgccccatacctgaagaccaaagttatctctgtgtgaacccaagacacgttcgagcaa	2355
QY	2228	TACCATTTACCTGCCGATGTCCCCCGCCACTTTAGATTATTTGATGCAAGTTTGGAAATTA	2287
Db	2354	taaccttgaacctgcgagatgtcccccgacatttagatltcatgtacgtatlttggaaataa	2413
QY	2288	CGGTAAAGGTCTGTAGCCGCTCAGCAGAGAGGCGAGTTTGAAGTGTGCTACCTTTGACATATGA	2347
Db	2414	tgggtaaaggatgcctgaacccctcagcagagagcgacgttttgaagttccctcaactttgacaatga	2473
QY	2348	TCCTGACCTGGAGTGTGTCTACCTCCCGCATGTGAGAGAGCTGAAGAACCAAGACTGC----	A 2403
Db	2474	gttgaaccttcggaatgtgcgtactccctcccatgtgaagatcgtgaagaaacgaaatctgcagaa	2533
QY	2404	GAGACGTCGACTTTGAGACACCCGCGCGGTGCTCCACCCCTAACAGCGCGGACCCCATATGC	2463
Db	2534	gatacgaactgaagcgcccaacctgaatctgtgcacccctcacaaagcaaacaccccaagatca	2593
QY	2464	TCTGAACCTCTTACCTTTGTGTGGTTCCAGATTTTGTTTTAAATTTTCTTACTTGTCTATTC	2523
Db	2594	tcgtgaactactaaacttltgtgtccagaa---tlttlttaatctccactcttgcatac	2650
QY	2524	TTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGAAGTGTGGGTGATTAACGT	2583
Db	2651	tttggcaactcgtgcactttaaataatagaataatgaatgtgaatgttgggtgtatctgtctt	2710
QY	2584	TATGTAA-----AGAGAGAGACCTGTGAGTCTGGGAGATGAGGCTGAGAGCAGAGAG	2635
Db	2711	tatctaataatgaataatagaatgtgtctctctgaagcccatgtatcagagagatgtgtggggggg	2770
QY	2636	GAGGCAAAAGGGA-----ACACTCTCTGTCTGCCCGCGCTGCGCTCTCTTTT	2682
Db	2771	gtaggtatagaagaaagaaagaaatgtctgttcttlttltgttccctgtccctcttct	2830
QY	2683	CAGCGACTCGGGGGT-----TGGTTGTAGACAGTGTGCTGCTGCTGCTCCATGCG	2732
Db	2831	cagcgacttlttgtatltgtltgtltgtltgttcttgaagaaatgtccctcctgtgtcgtgcgc	2890
QY	2733	TACTGTGTGCCCACTGTGTGAGCTGATACCCCAT-----CTGGCAACTCTGTGG	2782
Db	2891	atctcctgtcctgttcttctgaagaatgtccacagagccacatagctatacatatcctgtg	2950
QY	2783	CTGTGCACTTTC--AACCTGTGATATATCCACATAGAGTATAGACTA--AGCCACGAGAG	2839
Db	2951	catgtgacttcttaacctgtgtgcacccaataatagaatagtagactatacttaagccctagg	3010
QY	2840	TTTCTCTTTAATTTAAAAAATAAAAAA 2868	
Db	3011	tttcttlttaataatagaataataataa 3039	
RESULT 10			
ID AAA34938			
AAA34938 standard; DNA: 2847 BP.			
XX			

[illegible]

QY	2221	GCAGCAATCCATTGACCTGCCATGTCGCCCGCCGACTTTAGATTCAATTGATGCAATTCG	2280
Db	2433	gcagcaataacacattgcacctgcgcatgtcccccgcgcttagatccaattgatgcagcttg	2492
QY	2281	GAAGTAACGAGTGGAAGGTGCTGAGACCCCTCAGCAGAGAGGGCAGTTTGAGTCCGCTCACGTTTG	2340
Db	2493	gaataataigtgtaaggtgtcctgaaccctcagcagaaggcgagtttgagtcctccactcttg	2552
QY	2341	ACATGATCTGACCTCGGAATGTGCTACTCTCCCATGTGAGAGCTGAAGACCAAGACT	2400
Db	2553	acatgagattgacctcggagtgctgactacccctcccatgtgagagctgagaaeggaagct	2612
QY	2401	GC---AGAGACCTGACTTGAGACACCTGCGCCGCTGCTCCACCCCTAAAGCCGGAACCC	2456
Db	2613	gcagaaagaatacagcactgcgagcgccctactcgtactctgcacaccctcacaacagcaaaccc	2672
QY	2457	CATATCGCTCGAAGACCTCTTAACCTTTGTGTGTTCCAGATTTTCTTTTAATTTCTACTTC	2516
Db	2673	cagatcatctcgaacactaactactcttggttccaga---tttttttlaattcctactctc	2729
QY	2517	TGCTATCTTTGGGCAATCTGCGACCTTTTAAAGAGAGAAGTAAGTAGAGTGAGTGAGAT	2576
Db	2730	tgcatactcttgagcaactcggcgcaactcttaaaaaatagagaatcgtgaatctgagtgat	2789
QY	2577	AAACTGTATGTATAA 2591	
Db	2790	ctgctttatctataa 2804	
RESULT 11			
ID	AAAX29281	AAAX29281 standard; cDNA; 2344 BP.	
AC	AAAX29281;		
XX	11-JUN-1999	(first entry)	
DT			
XX	Human STAT3 allelic variant encoding cDNA.		
XX			
KW	Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;		
KW	intracellular, transcription factor; interleukin-6; medicament; variant;		
KW	pharmaceutical; autoimmune disease; inflammatory; human; ss.		
OS	Homo sapiens.		
XX			
PN	EP905234-A2.		
XX			
PD	31-MAR-1999.		
XX			
PF	18-FEB-1998: 98EP-0102774.		
PR	16-SEP-1997: 97EP-0116061.		
XX			
PA	(ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.		
XX			
PI	Della Pietra L, Serlupi-Crescenzi O;		
XX			
DR	WPI: 1999-192664/17.		
XX			
DR	P-PSDB: AAY03768.		
XX			
PT	New human Signal Transducer and Activator of Transcription 3 (STAT3)		
PT	allelic variant useful for treatment of autoimmune and inflammatory		
XX	disease		
XX			
PS	Claim 5; Page 9-13; 32pp; English.		
XX			
CC	The present sequence represents a cDNA encoding a predominant allelic		
CC	variant of human Signal Transducer and Activator of Transcription 3		
CC	(STAT3) protein, an intracellular transcription factor which mediates		
CC	IL-6 signals. This sequence differs from the original published human		
CC	STAT3 gene sequence (Akira et al., 1994). Host cells comprising a		
CC	recombinant STAT3 DNA molecule can be used for the recombinant expression		
CC	of the variant. STAT3 protein is useful as a medicament or pharmaceutical		

CC composition for treatment of autoimmune or inflammatory diseases.
XX Best Local Similarity 91.3%: Pred. No. 0:
SQ Sequence 2344 BP: 649 A; 580 C; 620 G; 495 T; 0 other:

Query Match 70.1%: Score 2012.6; DB 20; Length 2344;
Matches 2135: Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 69 ATGGCTCAGTGAACACCGTGCAGCTGACACACGCTTACCTGAAGCAGCTGCACG 128
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 atgcccattgatatcagctacagcagcttgacacacggtacctgagcagcctccacg 60
QY 129 CTGTACAGGACGCTTCCCACTGGAGTGGCGGAGTCTTCTGCGACCTGGACTTGGATTGAG 188
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ctctacaagtgcagcttcccacaatgagcctgagcagcttctgcccccttgattgagagt 120
QY 189 CAAGACTGGGCATATGTACAGCAGCAAGATGCATGACAGTGTGTTCTTAATCTTC 248
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 caagattgagcatatgagcagcagaagaatcacatgccccttgglttctataatctc 180
QY 249 TTGGGTGAATTTACACAGCAATATACCGATTCCTGCAAGATGCCAATGCTCTATCAG 308
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 ctgggagagattgaccagagatagcgccttcgcaagagtcgaatgctctatcag 240
QY 309 CACACCTTCTGAGAAATCAAGCAGTTTCTGACAGCAGATATCTTGAAGCCCATGGA 368
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 caaatctacagaagaatcgaagcagcttcttcagagcagagatcttgagaagccaatgag 300
QY 369 ATTGCCCGATCTGCGCCGATCCCTGTGGAAGAGTCTGCTCTCCACAGCCGACGC 428
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 atggcccgattgtagccggtgctgcttggaagaatcagcgtctctacagactgagcc 360
QY 429 ACGGACGCCAGCAAGGGGCGCCAGCAACCCACCAAGCCGCGAGTACAGAGAG 488
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 acgpgccagagaagggcgccagcccaaccacacagcagcgtgtagagagagag 420
QY 489 CAGCAGATGTTGAGCAGCAGCTTCTCAGAGATGTCGGAAGCGAGTGCAGATCAGAAC 548
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 cagcagatgctgagcagcagccttcagatgctcaggaagagtgagagctctagaacag 480
QY 549 AAAATGAAGTGTGTGGAACCTTCAGAGCAGTGTGATTTCAACTCAAAAACCCCTCAG 608
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 aaaaataaagtgtgtagaagatctccagatgacttgattcaactataaaacccctcaag 540
QY 609 AGCCAGGAGACATGCAGATGTGAATGGAACACACAGTCTGTGACACAGCAAGAGATG 668
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 agtcaagagacatgcaagatctgaaatgaaacaacacagtcagtgacagcgcaagatg 600
QY 669 CAGCAGCTGGAACAGATGCTCACAGCCCTGAGCCAGATGCGGAAGCATTTGAGTGA 728
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 cagcaactgtgaacagatgctcactgctgagaccagatgcygaagcactcgatgag 660
QY 729 CTGCGCGGCTCTTGTGACGATGAGTACGTCAGAAAGCACTGATCTGAAGACCTG 788
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 661 ctgagcgaggtcttctgcaagatgagtagctgcycaaaaaactctcacgagcgagagctg 720
QY 789 GCTGACTGAAGAGCGGCCAGATGCGGTGCATGCGAGCGCCCTCCCAACATCTGCTG 848
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 721 gctgactggaagagcgagcaacagatgctgcatctgagcccgcccaacatctgcta 780
QY 849 GACGCTGTGGAATAAGTGAATTAATTAGCAAGATCTCACTTGAAGCCGCCAACA 908
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 781 gatcgcgtctgaaaaactggaataacgtcatatagcagaatctcaactctgagcgcctcaaaa 840
QY 909 ATTAAAGAACTGGAGAGCTGCAGAGAAAGTGTCTTACAAAGGCGACCTATCTGTGAC 968
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 841 attaagaataactgagaggtctgcaaaaaagttctctacaagaagggaaccccatctgacag 900
QY 969 CACCGGCCCATGCTGAGAGAGAGAGATGCTGTGACACTGTTCAAGAACTTAATGAAGATGCC 1028
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 901 caccggtcgatgctgtagagagagaatcgtgagcgtgtttagaaaacttaatgaaaagtgc 960

QY 1029 TTGCTGTGAGCGCGGACGCTTCGATGCCATGACCCGAGCGGCTTTAGTCATCAAG 1088
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 ttctggtgtagagcgagcagccctgcatgcatctccgaccgagccctctgcatcaag 1020
QY 1089 ACTGATGTCAGTTTACACGAAAGTCAAGTGTGCTGTGCTCAAAATTTCTGATGTTGAATTA 1148
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1021 accggtgctcagctactactactaaagtcaggtctgctgctcaaatccctgagttgattat 1080
QY 1149 CAGCTTAAATTTAAATGTGTGATGTAATAAGACTGTGGGAGATGTTCTGCTCCCTCAGAGGG 1208
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1081 cagcttaaaatlaaagltgcatctgcaaaagactctgaggacgttgacgtctcagaagga 1140
QY 1209 TCTCGGAAATTTAATTTGTGGCAGCAACAAAAAGTATGAACATGAGAGAGTCTAAC 1268
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1141 tcccggaatcttaactctcttgagccacaacaacaagaatgataacttgagaatccaaac 1200
QY 1269 AACGGCAGCTGTCTGACAGATTCAAGCACCTGACCTTTAGGGACGAGAGATGTGGAA 1328
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1201 aacgcaagcctctctgcaaaatcacaacacttgaccctgagagagagagatgtaggaat 1260
QY 1329 GGAGGCGGTGCAATTTGTATGCTCTTGTGATGCTGACTGAGAGAGCTGCACCTGATCAC 1388
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1261 ggggagccgaagcaattgtagtcttccctgattgtagctggaagagctgcaactgacac 1320
QY 1389 TTGAGACTGAGTGTACCAAGGCTCAAGATTTGACCTTAGAGACCACTCTTGCCA 1448
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1321 ttgagacagcaggtgcatcacaagccccaagattgacctagaagcccaactcttgcca 1380
QY 1449 GTTGTGTGTATCTCCAAACTCTGTGATGCCAAATGCTTTGGGCAATCCTGTGTGAT 1508
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1381 gttgtgtgtatctcccaactctgcatagccaatgacctgagcgctgcatcctgtgtag 1440
QY 1509 AACATCTGACCAATTAACCCCAAGAAAGCTGAATCTTCTACTTAAGCGGCAATTTGGAAC 1568
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1441 aacatgtgtacccaacaaatcccaagaatgtlaaactcttlaaccaagcccccaattggaac 1500
QY 1569 TGGGACCAAGTGGCCAGAGTGTCTACGTGGCAGTCTCTGTCACACCAAGCAGGCGCTG 1628
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1501 tggatcaaaatgagcagaggtccctgagctgtagcagttctctcccccacaaagagagactg 1560
QY 1629 AGCATGAGCAGCTGAACAGCTGTGGCTGAGAACTCTAGGCGCTGCTGAACTCTCA 1688
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1561 agcatcgaagcagctgacataacatgagagaacactcttgagacctggtgtgaattctca 1620
QY 1689 GGGTGTGAGATCATGCGGCTTAATTTCTGCAAGAAAGCATGGCTGGCAAGGCTTCTCC 1748
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1621 ggggtgtcagatcagctgaggttaatttgcaagaagaacatgctgcaagggctctcc 1680
QY 1749 TTCTGGGTCTGGCTAGCAATATCATGACCTTGTGAAAAAGTATATCTTGGCCCTTTGG 1808
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1681 ttctggtgtcgtgtagacaatatcatctgacacttgtagaaaagatactctgagcccttgg 1740
QY 1809 AATGAAGGATACATCATGAGTGTATCATCAGCAAGAGCGGAGCGGCACTCTTAAGACA 1868
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1741 aaggaagagtaactatctatggtcttalcagtaaggaacgaggaagggcactcttgagcact 1800
QY 1869 AAGCCCGGGCACCTTCTACTGCGCTTTCAGCGAGCAGCAAGAGAGAGGGTCACT 1928
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1801 aagcctcccaagcactcttcctgctaagatctcagtgaaagcagcaagaagagagcgctacat 1860
QY 1929 TTTCATTTGGTGGAAAAAGCATCAGTGGCAAGACCCAGCTTCAGTCTTAGAGCCATTAC 1988
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1861 ttcaacttggtgtagaagagacacacagcgtgtaaaacccagatccagtcggtgaaacataac 1920
QY 1989 ACCAAGCAGCAGCTGAACAACATGTATTTGCTGAATCATGTGGCTTTAAGATCATG 2048
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1921 acaagacagcagctgtaacaacatgcatcttgctgaatacatcatctgggtctataagatcatg 1980
QY 2049 GATGCAACCAACTCTGTGTCTCCACTTTGTCTACTCTACCCGCAACTTCCCAAGAG 2108
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1981 gatgtactacaatatctctggtctcactcactgctatctatctctgacatctcccaagag 2040
QY 2109 GAGGCAATTTGAAAGTACTGTAGCCCGAGAGCCAGAGACACCCCGAAGCCGACAGGT 2168

Db 2041 gagcattcggaaagatattgctgcgacagagaccagagatcctgaagctgacccaagt 2100
 ||||||| ||||||| || ||| |||||||||||| || ||||| |||||||
 QY 2169 AGTGTGCGCCCGTACTGAGACCAAGTTCATCTGTGTGACACCAAGACCTGCACCAAT 2228
 || ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 2101 agcgtcgcccatcctgaagaccagtattatctgtgtgacacaaacgacgtgcaaat 2160
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 2229 ACCATTGACCTGCGATGTCTCCCGCCTTTTATGATTGATGCAAGTTTGAATAAC 2288
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 2161 accattgacctgcgaatgctcccccgaacttagattcattgacagtttgaaataat 2220
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 2289 GGTGAGAGTGTGAGCCCTCAGCAGAGGAGGACATTGAGTCGCTCAGCTTGTGACATGAT 2348
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 2221 ggtgaagagtgctgaaccctcagcagagagcagttgagtcctcctcattgacatggag 2280
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 2349 CTGACCTCGGAGTGTGTCTACCTCCCGCATGTGAGAGACCTGAACCAAGCTGCAGAGA 2407
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 2281 ttgacctcgagtgctgctacctcccccattgtgagagctgaaagcgtgcagaaa 2339
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 12

AAT31278
ID AAT31278 standard; cDNA: 2277 BP.

XX AC AAT31278:

XX DT 24-OCT-1996 (first entry)

XX DE Mouse STAT1 cDNA.

XX KW STAT1: signal transducer and activator of transcription;

KW STAT1: DNA binding protein; ligand: receptor; oncogenesis;

KW Inflammation; autoimmune disease; antagonist; therapy; ds.

XX OS Mus sp.

XX FH Key Location/Qualifiers

XX FT CDS 4..2254

XX FT /*tag- a

XX PN WO9620954-A2.

XX PD 11-JUL-1996.

XX PF 28-DEC-1995; 95MO-US17025.

XX PR 06-JAN-1995; 95US-0369796.

XX PA (UYRQ) UNIV ROCKEFELLER.

XX PI Darneil JE, Horvath CM, Wen Z, Zhong Z;

XX DR WPI; 1996-33941/33.

XX DR P-PDB; AAM03172.

XX PT New STAT protein DNA-binding domain peptide(s) - useful for

XX CC oncoanalysis, preventing or treating cellular dysfunction, e.g.

XX CC PT oncogenesis, inflammation, parasitic disease or autoimmunity

XX CC PS Disclosure; Page 82-85; 138pp: English.

XX CC A cDNA clone (AAT31278) codes for mouse signal transducer and

XX CC activator of transcription (STAT) protein STAT1 (AAM03172).

XX CC a homologue of human STAT1-alpha (AAM03168), that serves a dual

XX CC purpose, i.e. signal transduction from ligand-activated receptor

XX CC kinase complexes followed by nuclear translocation and DNA binding

XX CC to activate transcription. STAT1 includes a DNA-binding domain

XX CC useful for screening for antagonists that modulate STAT activity.

XX CC The cDNA can be used for prohn. of STAT1 or mutant STAT1.

XX CC Sequence 2277 BP; 661 A; 555 C; 562 G; 499 T; 0 other;

Query Match 23.8%; Score 682; DB 17; Length 2277;
 Best Local Similarity 59.9%; Pred. No. 76-180;
 Matches 1239; Conservative 0; Mismatches 805; Indels 24; Gaps 5;
 QY 65 CAGATGCGCTGAGAACACAGCTGCAGCAGCTGCAGACAGCCTACCTTACAGCAGCTGCA 124
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 1 caggatgacagtggttcgagcttcagcagctgcagctcccaagcttcgagcaggtccca 60
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 125 CCACTGTACAGCGACACGCTTCCCATGAGAGCTGCGGACGTTCTTGACCTTGAATTA 184
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 61 ccagctgtaacagagttcccatggaatacagacagtaacctggtcccgagtcgtga 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 185 GACTCAAGAGTGGCATATGACACCCAGCAAGAGTACATGCCACCTGTGTTTCATTA 244
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 121 aaagcaagacttggagcaacgctgctcattatgcttggttgacacatccgctccatga 180
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 245 TCTCTTGGGTGAATTTGACCAAGCAATATAGCCGATTCCTGCAAGAGTCCAATGCTCTTA 304
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 181 cctcctcagctgagcagaccagacagccgcttctctctgaggaataattcttgct 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 305 TCAGCAACACCTTGAAGATCAAGCAGTTTCTGCAGACAGGATCTTGAGAGCCCAAT 364
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 241 gcagcaacaatacgaagaagcagctcattatgcttggttgacacatccgctccatga 300
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 365 GGAATTTGCCCGGATCGTGCCTGATGCTGTGGAAGAGTCTGCTCCAGACGCGC 424
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 301 acagatgctcatgcatcattcaacactgctcgaagaagaagaatcttggaataatgc 360
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 425 AGCCACGCGCCAGCCAGGCGGCCAGCCACACCCACACGCGCGGTAGTGACAGA 484
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 361 ccaaatattatca-----ggccgaagaggaataatcagaacaactgtagttaga 414
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 485 GAAGCAGCAGATTTGTGAGCAGCATCTTCAGATGTCCGGAAGCGTGAGATGTAGA 544
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 415 taacagaaagagctgacagtaagtcagaatgtaagtaagtaagtaagtaagtaagta 474
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 545 ACAAGAAATGAGAGTGTGAGAGAACTCCAGAGCAGCTTGTGATTACACAAACCTT 604
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 475 gcaggaatcaaaccttaagaataatcaagaatgaatgaatgaatgaatgaatgaatga 534
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 605 CAAGAGCCAGAGACATGACGATGTGAATGGAACCAACCACTGTGTGACAGCAGAA 664
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 535 tcgaagaagaagaag-----gtgaagcacaatgltggtgaggaagagaccacaagaaga 588
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 665 GATGACAGCCTGGAACAGATGCTCACAGCCCTGGAACAGATCCGGAAGCAATTGTAG 724
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 589 acagctgctgctcacaagaatglttlaatgcttgacaataagaagaagaagaataatca 648
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 725 TGAGCTGCGCGGCTCTTGTACAGCAATGAGTACGTGCAGACAGCAGCTGATGAGA 784
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 649 caaatcagaagtlgtctgaattccatcgagctcaatcaagaacacatcgtatgaagga 708
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 785 GCTGGCTGACTGAGAGAGCGCCAGAGATGCGTGCATGAGAGCCCTCCCAACTCTG 844
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 709 gctcgtgagtgtaagaagaagcagcagcgtcgtgattcgtggggagccgcaagcgtg 768
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 845 CTTGAGCCGCTGTGAAATCTGATTAATTATGACAGAAATCTCACTTATGAGCCGCCA 904
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 769 cctgatacagctcaaacgctgctcaccatctgttcagagagaccctcagcagatccgta 828
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 905 ACAATTAAGAACTGAGAGGCTGCAGCAAGAAAGTCTCTACAGGCGCACCTTATCGT 964
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 829 gcaacttaaaagctcgaagagagtgtaacaagaatcactatgaagccgacctattac 888
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 965 GCACACCGCCCATGCTGAGAGAGAGATGCTGAGAGCTTGTCAAGAACTTAATGAAGAG 1024
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 889 aaaaaacaagcagtgltgctcagatcgaacactccctcctccacagagctcatcagag 948
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 1025 TGCTTGTGTGAGAGCGGACGCTTCATGCGCATGACCGGACCGGCGCTTAGTCAT 1084
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 949 ctccttggtgtaagaagaagcgcgtgcatcaaacacacccgcaagagccctggtctt 1008
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 QY 1085 CAAGACTGTGTCCAGTTTACCAAGCAAGTCAGGTTGCTGTCATAATTTCTGAGTTGAA 1144

```
Db 1009 gaagactggggtacagctcactcgtcgaagtcgagactggtgtgtaaaatctgcaagagtcgaa 1068
1145 TTATCAGCTTAAATTAAGGTGTGATGTGATAAGACTCTGGGGATGTTCTCTCCCTCAG 1204
Db 1069 tctataacgaagaagtgaaatgtaactcttgacaagaatgtaacgaagaaacacagttaa 1128
1205 AGGCTCTCGAAATTTAATCTTGGGCAACGAAACAAAGATGATGAATGAGAGAGTGC 1264
1129 aggatttcggaagtccaacatcttgggtacgacacaaaagtatgaacttgaagaatc 1188
1265 TAACACGCGACGCTGTCTGACAGATTCAAGCACTGACCTTAAAGGAGCAGAGATGTGG 1324
Db 1189 caaccaaggaagctcgtcgaagtcgacccacacactgcaacttgaaggaacaagaaa---- 1243
1325 GAATGAGGCGCTGCCAATTTGATGCTCTTCACTGATGCTGACGTGAGAGAGTGCACCTGAT 1384
Db 1244 -aagctctgggaaacagaactaaatgaagggtccctcactatgttcaaccgaagaactcactc 1302
1385 CACCTTCGAGACTGAGGTGTACCCAGCAGGCTCAAGATTGACCTAGAGACCTCTCTT 1444
Db 1303 tagcttgaaccagctgtgacagccagctgtgattgactctgagagccactctc 1362
1445 GCCAGTTGTGTATCTCCACATCTGTGATGCCAAATGCTTGGGATCAATCTCTGTG 1504
Db 1363 tccctgtcgtgtatctccaacgtcagccagctccccaagctggtcgtctactcgtg 1422
1505 GTATAACATGCTGACCAATTAACCCCAAGAACGTGAACTTCTCACTAAGCGCGCAATTGG 1564
Db 1423 gtacaacaagctgtgtaacagagcccaagaaatcctcctcttccgtaaacccccgtgcgc 1482
1565 AACCTGGGACCAAGTGGCCGAGGTGCTCAGCTGGGAGTTCTCGTCCACCAAGCAGAG 1624
Db 1483 tgggtgtcccaagctcgaaggtgtgtgagtggtgcaatttcaatcgaacgaagag 1542
1625 GCTGAGCATCGAGCAGCTGATCAGACGCTGTGTAAGCTCTAGGGCTGTGTGAATGA 1684
Db 1543 tctgaacgacagacactgagcagctgtggaagaagctgtcgtggtccctaaatgc---tgg 1599
1685 CTCAGGCTGTGACATCATATGAGCTTAATTTCTCAAAAGAAACATGCTGGCAGAGCTT 1744
Db 1600 cctcgtgtctatctcactcgaagaggttctgaagaaataataatgaataaattt 1659
1745 CTCCTTCTGGTGTGCTAGACATATCATGACCTTGTGAAAAGTATATCTTGGCCCTT 1804
Db 1660 tctcttctgtccttgtatgtacacacatcctlagagctcatalaagaagacctgtgtgcct 1719
1805 TTGGAATGAAGGTATCATGTGTTTTCATGACGAGGAGCGGAGCGGCGCATCTCTAAG 1864
Db 1720 ctggaatgtatgtgtatcatatgtgtcctcactcgaagaagagcgaagcgcctcgtcaa 1779
1865 CACAAAGCCTCCGGGACCTTCTTACTGTGCGTTTCAAGCAGAGCAGCAAGAGAGGGGT 1924
Db 1780 ggaacagcagcaggaagctcctgtctagattcagtgagagctcccggaagggccat 1839
1925 CACTTTCATTTGGTGTGAAAAGAGC---ATCAGTGGCAGAGCACCAGATCAGCTCTAGA 1981
Db 1840 ccaatcacaatgtgtggaacggtcccaagaacggtgaacactgactcatalgtcgtg 1899
1982 GCCATACACAGCAGCAGCAGCTGAAACATGTCTATTTGCTGAATCATGAGCGCTATTA 2041
Db 1900 gccctcaagaagaaagaacttccagctgttacttccagatatatttcgcaactcaaa 1959
2042 GATCATGATGCGACCAACATCTGTGTCTCCACTTGTCTACCTCTACCCGACATTC 2101
Db 1960 agtcatgtgtcgtggaataatcaccagaagatcccccgaagtatctgtaccccaatttga 2019
2102 CAAGGAGGAGGATTTGGAATGACTGT 2129
Db 2020 caaagaccagccttgggaagatatt 2047
```

RESULT 13

```
AA089338
ID AA089338 standard; cDNA; 2277 BP.
XX
AC AA089338:
XX
DT 27-SEP-1995 (first entry)
XX
DE Mouse Stat1 cDNA.
XX
KW Signal transducer and activator of transcription; STAT; Stat1;
KW Stat1; receptor recognition factor; transcription factor;
KW cellular denaturation; denaturation; dysfunction;
KW interferon-gamma; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 5..2254
FT /tag= a
FT
MO508629-A.
PN
XX 30-MAR-1995.
PD
XX
PF 26-SEP-1994; 94MO-US10849.
XX
PR 24-SEP-1993; 93US-0126588.
PR 24-SEP-1993; 93US-0126595.
PR 11-MAR-1994; 94US-0212184.
PR 11-MAR-1994; 94US-0212185.
XX
PA (UVRQ ) UNIV ROCKEFELLER.
XX
PI Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;
XX
DR WPI; 1995-139598/18.
XX
DR P-PSDB; AAR72080.
XX
PT Receptor recognition factor implicated in transcriptional
PT stimulation of genes - useful in drug screening assays and/or
PT for treating cellular denaturation, denaturation and/or
PT dysfunction, etc.
XX
PS Claim 10; Page 96-99; 160pp; English.
XX
CC A fragment encoding the human Stat1 protein was used to screen a
CC murine thymus and spleen cDNA for homologous proteins. A highly
CC homologous gene (given in AA089338) was isolated that encoded a
CC 91 kDa protein (AAR72080) (Stat1) that was responsive to
CC interferon-gamma.
XX
SO Sequence 2277 BP; 661 A; 554 C; 562 G; 500 T; 0 other;

Query Match 23.7%; Score 680.4; DB 16; Length 2277;
Best Local Similarity 59.9%; Pred. No. 2e-179;
Matches 1238; Conservative 0; Mismatches 806; Indels 24; Gaps 5;

Db 65 CAGGATGGCTCAGTGAACAGCAGCTGACGAGCAGCTGACACACGCTACTGGAAGAGCTGCA 124
1 caggaatgacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 60
125 CCAGCTGTACAGCAGACAGTCCCATGTGAGTGTGCGGACGTTCTGACACTTGGATTGA 184
61 ccagctgtacagcagcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
185 GAGTCAAGACTGGGCAATATGACGACGCAAGCAAGATCATGATGCAAGTGTGTGTGATA 244
121 aagcaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
245 TCTCTGGGTGAATATGACGACGATATGAGCGATTCCTGACAGAGTCCATGCTCTGA 304
181 cctcctcacaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
```


PN W0200055320-A1.
XX 21-SEP-2000.
XX 08-MAR-2000: 2000MO-US05989.
PF 12-MAR-1999: 9905-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM.
XX WPI: 2000-579444/54.
DR P-PSDB: AAB54098.
XX
PT New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition -
XX
XX Claim 1: Page 560-561; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytosolic,
CC neuroprotective, nontoxic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC98232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention.
XX
XX Sequence 2974 BP: 835 A: 660 C: 741 G: 725 T: 13 other:
SO

Db 528 acagcatacctaagagaagcaagcgttaattcttcagataatttcaaggaagccaat 587
Qy 365 GGAATTTGCCCGATCGTGGCCCGATGCTGTGGGAAGTCTCGCTCCTCCAGACGC 424
Db 588 ccagatgctatagatcattcacaagctgctgaaagaagaagaaatctcgaagaaagc 647
Qy 425 AGCCACGGCACCACGACGAAGGGGCCAGUCCAACCCACCAACCCCGCTAGACAGA 484
Db 648 ccagagatllaatcaagcctcagctcggaagaaatllcaagagca-----glatgltaga 701
Qy 485 GAAGCAGCAGATGTTGGAGCAGATCTTCAGAGATGTCGGAAGCGAGTGCAGATTAGA 544
Db 702 caaagagagagagcttgacagtgaaagtgaaagatgaaagagagagtgatgltataga 761
Qy 545 ACAGAAATGAAGTGTGGAGAACTTCACAGACGACTTTGATTCAACTACAAACCT 604
Db 762 gcatgaatlcagagccctggaagatllcaagaagatllagacllcaaaagcaaacctc 821
Qy 605 CAAGACCCAGAGAGACATGCGATCTGAATGGAACAAACCTGCTGACACAGCAAA 664
Db 822 gcaagatcaagaaac-----gagaccaatggtgcaagagtgatcagaacaaga 875
Qy 665 GATGACGAGCTGGAACAGATGCTCACAGCCCTGGACCATGCGAGAACATTTGAG 724
Db 876 acagctttactcaagaagatgltatllatgcttgacaaatlaagaaagaaagtagtca 935
Qy 725 TGAGCTGGCGGGCTCTTCAGCAATGATGATGATGATGATGATGATGATGATGATG 784
Db 936 caaaatuaatagatgctgcaatgctgcaatgctgcaatgctgcaatgctgcaatgctg 995
Qy 785 GCTGCTACTGGAAGAGAGCGCGGAGAGATGCGTGCATGCGAGCCCTCCCAACATCTG 844
Db 996 actagtag 1055
Qy 845 CCTGACCGCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 904
Db 1056 cttagatcaagctgcaag 1115
Qy 905 ACAAAATTAAGAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 964
Db 1116 gcaagcttaaaagtgag 1175
Qy 965 GCAGACCGCGCCATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1024
Db 1176 aaaaacaacaagtgatgagagagagagagagagagagagagagagagagagagag 1235
Qy 1025 TGCTTGTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1084
Db 1236 ctgcttltgtag 1295
Qy 1085 CAAGACTGCTGCTCAAGTTTACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1144
Db 1296 gaagacag 1355
Qy 1145 TTATAGCTTAAATTAAGAGTGCATGATGATGATGATGATGATGATGATGATGAT 1204
Db 1356 ttaataattgaag 1415
Qy 1205 AGGCTCTGGAATTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1264
Db 1416 aggatttagaag 1475
Qy 1265 TAAGACGAG 1324
Db 1476 caacaaag 1500
Qy 1325 GAATGAT 1384
Db 1531 -aatgctgacacaaag 1589
Qy 1385 CACTTGGAT 1444
Db 1590 tagtttgaacccaatgtagcagcagcgtgttgaatgtagcagcagcagcagcagc 1649

QY	1445	GCCATGTTGTGTCATCTCCACATCTGTAGATGCCAAATGCTTGGCATCAATCCTCTG	1504
Db	1650	gcccgcttggtagtcatctccaacgtccagcagctcccgagcggttggcctccactcttg	1709
QY	1505	GTATTAACATCTGTAACCAATTAACCCCAAGCACTGAACATTTCTTCACTTAAGCCGCCAATTTG	1564
Db	1710	gtacaacaatcgtcgtggtcggaatacccgagactctgtctcttctccgactccacatgtc	1769
QY	1565	AACCTGGGACCAAGTGGCCGAGGTGCTCAAGTGGCAATTTCTGTCCACCAACCAAGCCAG	1624
Db	1770	agcagtgggcctcagccttccagaagtgctgagtgtagcagtttcttctccgacccaatgtc	1829
QY	1625	GCTGAGCATGAGCAGCTGTACAACGCTGGCTGGAAGCTCCTAGGCGCTGTGTTGAACCTA	1684
Db	1830	tctcaatgtggaaccaactgaactgtgtggagagagaagctctctgtctctaaagccagcc	1889
QY	1685	CTCAGGCTGTCAAGTACATGAGGCTAAATTTCTGCAAAAGAAACATGCTGGCAAGGCTT	1744
Db	1890	cgatg---gtctcatctccgtggagcaggtttgttaagaaaaataaatagtataaaat	1946
QY	1745	CTCCTTCGTGGTGGCTAGACATATCATCGACCTTGTGAAGAAAGTATATCTTGGGCTT	1804
Db	1947	tccctcttggtcttgatgtgaagcaactagaacccaataaaacactgtctccct	2006
QY	1805	TTTGAAATGAAGGTATCATATGAGTTTCATCAGCAAGAGAGCGGAGCCATCTTAAG	1864
Db	2007	ctggaatgtagtgcatactatggcttcaatcagcaagagcgtagcggttctgttga	2066
QY	1865	CACAAAGCCCCGGGCGACCTTCTACTGCGGCTTCACGCGAGACAGAAAGAAAGAGGAGGT	1924
Db	2067	ggaaccagacccggagacctctccgtctggtctcagtgagaggtcccggaaggagccat	2126
QY	1925	CACTTTCACTTGGGTGCAAAAGGAC--ATCAGTGGCAAGACCCAGATCCAGTCTTGA	1981
Db	2127	ccatctcatcaggtgtgagcggtctccgaacgagagcgaaactgtactatgctgtga	2186
QY	1982	GCCATACCAAGCAGCAGCTGAACACATGTCAATTTGCTGAATATCATGGGCTATTA	2041
Db	2187	acctcaacagaagaagaacttctgtctgtactctccctgaatcatctcgaattaca	2246
QY	2042	GATCATGATCGCAGCAACATCTGATGTCTCCACTTGTCTACTCTTACCCTGACATTC	2101
Db	2247	agtcatagtctgtagaatatctctcgtgaatcccccgaagtagtctgtatccaatattga	2306
QY	2102	CAAGAGAGAGCATTTGGAAGTACT 2127	
Db	2307	caaaagacatgccttggaaaagtatt 2332	
RESULT 15			
AAQ89337			
ID	AAQ89337	standard; cDNA; 2607 BP.	
XX	AAQ89337:		
AC			
XX	27-SEP-1995 (first entry)		
DT			
XX			
DE	Human Stat84 cDNA.		
XX			
KM	Signal transducer and activator of transcription; ISGF-3; STAT;		
KM	Stat84; receptor recognition factor; transcription factor;		
KW	cellular debilitation; derangement; dysfunction;		
KW	interferon-gamma; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key location/Qualifiers		
FT	197..2335		
FT	CDS		
XX	/*tag= a		
PN	W09508629-A.		
XX			

PD 30-MAR-1995.
 XX
 PF 26-SEP-1994; 94WO-US10849.
 XX
 PR 24-SEP-1993; 93US-0126588.
 PR 24-SEP-1993; 93US-0126595.
 PR 11-MAR-1994; 94US-0212184.
 PR 11-MAR-1994; 94US-0212185.
 XX
 PA (UYRO) UNIV ROCKEFELLER.
 XX
 PI Darrell JE, Schindler CW, Shuai K, Wen Z, Zhong Z:
 XX
 DR MPI. 1995-139598/18.
 DR P-PSDB: AAR72079.
 XX
 XX Receptor recognition factor implicated in transcriptional
 PT stimulation of genes - useful in drug screening assays and/or
 PT for treating cellular debilitations, derangements and/or
 PT dysfunctions, etc.
 XX
 PS Disclosure: Page 91-94; 160pp; English.

AA The sequences of cDNA encoding receptor recognition factors having
CC
CC mol. wt. of 113 kDa (Stat113), 91 kDa (Stat91) and 84 kDa (Stat84)
CC are given in AA089335-37 and the deduced amino acid sequences of the
CC STAT proteins in AAR72077-79. These ISGF-3-derived proteins are
CC activated by binding of interferon-alpha (all 3 Stat proteins) or
CC interferon-gamma (Stat91) to cell receptors.
XX
SQ Sequence 2607 BP: 772 A; 566 C; 619 G; 650 T; 0 other:
XX

Sequence 2607 BP; 772 A; 566 C; 619 G; 650 T; 0 other;

Query Match	23.48;	Score 672.6;	DB 16;	Length 2607;
Best Local Similarity	59.28;	Pred. No. 3.2e-177;		
Matches 1250; Conservative	0;	Mismatches 839;	Indels 24;	Gaps 5

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2002, 14:00:58 ; Search time 63.17 Seconds
(without alignments)
10285,980 Million cell updates/sec

Title: US-08-212-185-11

Perfect score: 2869
Sequence: 1 GCCGCCGACGCCAGCCGCCG.....AATTAATAAAAAAAAAAAAA 2869

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA :
1 : /cgn2_6/ptodata/2/1na/5A.COMB.seq : *
2 : /cgn2_6/ptodata/2/1na/5B.COMB.seq : *
3 : /cgn2_6/ptodata/2/1na/6A.COMB.seq : *
4 : /cgn2_6/ptodata/2/1na/6B.COMB.seq : *
5 : /cgn2_6/ptodata/2/1na/PTOS.COMB.seq : *
6 : /cgn2_6/ptodata/2/1na/Dackfile1.seq : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2869	100.0	2869	1	US-08-369-796-11 Sequence 11, Appl
2	2869	100.0	2869	2	US-08-852-091-11 Sequence 11, Appl
3	2869	100.0	2869	2	US-08-820-754-11 Sequence 11, Appl
4	2869	100.0	2869	3	US-08-956-652-11 Sequence 11, Appl
5	2869	100.0	2869	3	US-08-956-869-11 Sequence 11, Appl
6	2869	100.0	2869	3	US-08-948-547-11 Sequence 11, Appl
7	2869	100.0	2869	5	PCT-US95-17025-11 Sequence 11, Appl
8	2864.2	99.8	2869	3	US-09-288-461-82 Sequence 82, Appl
9	2861	99.7	2869	4	US-09-364-970-9 Sequence 9, Appl
10	2415.6	84.2	2852	1	US-08-416-581B-7 Sequence 7, Appl
11	2415.6	84.2	2852	1	US-08-416-581B-8 Sequence 8, Appl
12	2294	80.0	2310	1	US-08-416-581B-6 Sequence 6, Appl
13	2144.2	74.7	2787	1	US-08-416-581B-3 Sequence 3, Appl
14	2144.2	74.7	2787	1	US-08-416-581B-4 Sequence 4, Appl
15	2144.2	74.7	2787	3	US-09-288-461-1 Sequence 1, Appl
16	2144.2	74.7	2787	4	US-09-087-465-5 Sequence 5, Appl
17	2026.8	70.6	2310	1	US-08-416-581B-2 Sequence 2, Appl
18	682	23.8	2277	2	US-08-369-796-7 Sequence 7, Appl
19	682	23.8	2277	2	US-08-852-091-7 Sequence 7, Appl
20	682	23.8	2277	3	US-08-820-754-7 Sequence 7, Appl
21	682	23.8	2277	3	US-08-956-652-7 Sequence 7, Appl
22	682	23.8	2277	3	US-08-956-869-7 Sequence 7, Appl
23	682	23.8	2277	3	US-08-948-547-7 Sequence 7, Appl
24	682	23.8	2277	5	PCT-US95-17025-7 Sequence 7, Appl
25	672.6	23.4	2607	2	US-08-369-796-5 Sequence 5, Appl
26	672.6	23.4	2607	2	US-08-852-091-5 Sequence 5, Appl
27	672.6	23.4	2607	2	US-08-820-754-5 Sequence 5, Appl

ALIGNMENTS

28	672.6	23.4	2607	3	US-08-956-652-5	Sequence 5, Appl
29	672.6	23.4	2607	3	US-08-956-869-5	Sequence 5, Appl
30	672.6	23.4	2607	3	US-08-948-547-5	Sequence 5, Appl
31	672.6	23.4	2607	5	PCT-US95-17025-5	Sequence 5, Appl
32	672.6	23.4	3943	1	US-08-369-796-3	Sequence 3, Appl
33	672.6	23.4	3943	2	US-08-852-091-3	Sequence 3, Appl
34	672.6	23.4	3943	2	US-08-820-754-3	Sequence 3, Appl
35	672.6	23.4	3943	3	US-08-956-652-3	Sequence 3, Appl
36	672.6	23.4	3943	3	US-08-956-869-3	Sequence 3, Appl
37	672.6	23.4	3943	3	US-08-948-547-3	Sequence 3, Appl
38	672.6	23.4	4003	5	PCT-US95-17025-3	Sequence 3, Appl
39	672.6	23.4	4003	4	US-09-087-465-1	Sequence 1, Appl
40	666.2	23.2	3943	4	US-09-364-970-10	Sequence 10, Appl
41	564.4	19.7	2375	1	US-08-369-796-9	Sequence 9, Appl
42	564.4	19.7	2375	2	US-08-852-091-9	Sequence 9, Appl
43	564.4	19.7	2375	2	US-08-820-754-9	Sequence 9, Appl
44	564.4	19.7	2375	3	US-08-956-652-9	Sequence 9, Appl
45	564.4	19.7	2375	3	US-08-956-869-9	Sequence 9, Appl

RESULT 1

US-08-369-796-11
Sequence 11, Application US/08369796

Patent No. 5716622

GENERAL INFORMATION:
APPLICANT: James E. Darnell, Jr.
APPLICANT: Zilong Wen
APPLICANT: Curt M. Horvath
APPLICANT: Zhong Zhong

TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:
ADDRESS: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey

COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/369,796

FILING DATE: 06-JAN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET INFORMATION: 600-1-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2869 base pairs

TYPE: nucleic acid
STRANDEDNESS: both
MOLECULE TYPE: CDNA

HYPOTHEICAL: NO
ANTI-SENSE: NO

ORIGINAL SOURCE:
ORGANISM: Mouse

IMMEDIATE SOURCE:
LIBRARY: splenic/lymnic

CLONE: Murine 195f6

FEATURE:
NAME/KEY: CDS
LOCATION: 69..2378
us-08-369-796-11

Query Match 100.0%; Score 2869; DB 1; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGACGACGAGCGCGCGAGTCGCGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT 60
DB 1 GCGCGACGACGAGCGCGCGCGAGTCGCGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT 60
QY 61 GCAGCAGATGCTCAGTGGACACAGCTGCAGACGTGACACACCTACCTAGAGAC 120
DB 61 GCAGCAGATGCTCAGTGGACACAGCTGCAGACGTGACACACCTACCTAGAGAC 120
QY 121 TGCACGAGTGTACAGCAGCAGTCCCATGAGCTGCGGAGATTCTGGACCTTGA 180
DB 121 TGCACGAGTGTACAGCAGCAGTCCCATGAGCTGCGGAGATTCTGGACCTTGA 180
QY 181 TTGAGAGTCAAGACTGGCATATGACAGCAAGAGTCAATGCCAGCTGTTTC 240
DB 181 TTGAGAGTCAAGACTGGCATATGACAGCAAGAGTCAATGCCAGCTGTTTC 240
QY 241 ATATATCTTGGGTGAATTTGACAGCAATATACCCGATTCTGCAAGAGTCCAAATGCC 300
DB 241 ATATATCTTGGGTGAATTTGACAGCAATATACCCGATTCTGCAAGAGTCCAAATGCC 300
QY 301 TCTATCAGCACAACCTTGAAGATCAAGAGTTTTCAGAGCAGGTATCTTGAGAGC 360
DB 301 TCTATCAGCACAACCTTGAAGATCAAGAGTTTTCAGAGCAGGTATCTTGAGAGC 360
QY 361 CAATGGAATTTGGCGGATGTCGCGGATGCTGCGGAAAGATGTCGCTCTCCAGA 420
DB 361 CAATGGAATTTGGCGGATGTCGCGGATGCTGCGGAAAGATGTCGCTCTCCAGA 420
QY 421 CGGAGCGACGCGAGCCAGCAAGGGGGCCAGGCCAACCAACGCGCCGTAGTGA 480
DB 421 CGGAGCGACGCGAGCCAGCAAGGGGGCCAGGCCAACCAACGCGCCGTAGTGA 480
QY 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTCAGATGTCGCGAAGCAGATGAGATC 540
DB 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTCAGATGTCGCGAAGCAGATGAGATC 540
QY 541 TAGAACAGAAAATGAAGTGTGAGAACTCCAGAGCATTTGATTCAACTACAAA 600
DB 541 TAGAACAGAAAATGAAGTGTGAGAACTCCAGAGCATTTGATTCAACTACAAA 600
QY 601 CCTCAAGAGCAGAGAGACATGAGATGGAATGGAACACCACTCTTGACAGAC 660
DB 601 CCTCAAGAGCAGAGAGACATGAGATGGAATGGAACACCACTCTTGACAGAC 660
QY 661 AGAAGATGACAGCAGCTGAGAACAGATGCTCAGAGCCCTGAGCAATGCGGAAAGCAT 720
DB 661 AGAAGATGACAGCAGCTGAGAACAGATGCTCAGAGCCCTGAGCAATGCGGAAAGCAT 720
QY 721 TGAATGAGCTGCGGGGCTCTTGTACCAATGAGTACGTGAGAGACACTGACTATG 780
DB 721 TGAATGAGCTGCGGGGCTCTTGTACCAATGAGTACGTGAGAGACACTGACTATG 780
QY 781 AAGAGCTGGCTGACTGGAAGGCGCGAGAGATGCGTGCATGGAGGCCCTCCAAACA 840
DB 781 AAGAGCTGGCTGACTGGAAGGCGCGAGAGATGCGTGCATGGAGGCCCTCCAAACA 840
QY 841 TCTGCTGAGACCTCTGGAAGAACTGATTAATTCATTAGCAGATTCGAATTCAGACC 900
DB 841 TCTGCTGAGACCTCTGGAAGAACTGATTAATTCATTAGCAGATTCGAATTCAGACC 900
QY 901 GCCAACAATTTAGAAACTGAGAGCTGCAGCAGAAAGTGTCTTACAAAGGCGACCTTA 960
DB 901 GCCAACAATTTAGAAACTGAGAGCTGCAGCAGAAAGTGTCTTACAAAGGCGACCTTA 960

QY 961 TCGTGCAGCACCAGCCCATGCTGGAGAGACGATCGTGAAGCTGTTCAGAACTTAATGA 1020
DB 961 TCGTGCAGCACCAGCCCATGCTGGAGAGACGATCGTGAAGCTGTTCAGAACTTAATGA 1020
QY 1021 AGAGTGCCTTCGTGTGAGAGCGGCGGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1080
DB 1021 AGAGTGCCTTCGTGTGAGAGCGGCGGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1080
QY 1081 TCATCAAGAGCTGTGCTACATTTACAGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1081 TCATCAAGAGCTGTGCTACATTTACAGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 TGAATTAATGCTTAAATTAAGTGTGATTAAGATCTGGGAGATGTTGCTGCC 1200
DB 1141 TGAATTAATGCTTAAATTAAGTGTGATTAAGATCTGGGAGATGTTGCTGCC 1200
QY 1201 TCAGAGGCTCTCGAAATTTAATTTACATTTCTGGGACGAAACACAAAGTGTATGAGATG 1260
DB 1201 TCAGAGGCTCTCGAAATTTAATTTACATTTCTGGGACGAAACACAAAGTGTATGAGATG 1260
QY 1261 AGTCTAACACGCGAGCTGTGAGAGTTCAGAGTTCAGAGCTGAGGAGCAGAGAT 1320
DB 1261 AGTCTAACACGCGAGCTGTGAGAGTTCAGAGTTCAGAGCTGAGGAGCAGAGAT 1320
QY 1321 GTGGGAATGAGAGCGGCGCAATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATG 1380
DB 1321 GTGGGAATGAGAGCGGCGCAATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATG 1380
QY 1381 TGATCAGCTTCAGAGCTGAGGTGATGACCAAGAGCTCAAGATTTGACTGAGACCCACT 1440
DB 1381 TGATCAGCTTCAGAGCTGAGGTGATGACCAAGAGCTCAAGATTTGACTGAGACCCACT 1440
QY 1441 CCTTGCAGTTGGTGTGATGCTGCAACATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1500
DB 1441 CCTTGCAGTTGGTGTGATGCTGCAACATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1500
QY 1501 TGTGTATACATGCTGAGCAATTAACCCAGAGAGCTGAACTTCTTCACTAAGCCGCAA 1560
DB 1501 TGTGTATACATGCTGAGCAATTAACCCAGAGAGCTGAACTTCTTCACTAAGCCGCAA 1560
QY 1561 TTGGAACCTGGGAGCAAGTGGCGAGGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
DB 1561 TTGGAACCTGGGAGCAAGTGGCGAGGTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
QY 1621 GAGGCTTGAGCATGACACAGCTGACCAAGCTGCTGAGAACTTCTGAGGCTGTGTGA 1680
DB 1621 GAGGCTTGAGCATGACACAGCTGACCAAGCTGCTGAGAACTTCTGAGGCTGTGTGA 1680
QY 1681 ACTACTCAGGCTGTACATGATGAGGCTAAATTTCTGCAAGAAACATATGCTGCAAGG 1740
DB 1681 ACTACTCAGGCTGTACATGATGAGGCTAAATTTCTGCAAGAAACATATGCTGCAAGG 1740
QY 1741 GCTTCTCTTCTGGGCTGTGCTAGACAAATATCATGCACTTGTGAAATAATATATCTTGG 1800
DB 1741 GCTTCTCTTCTGGGCTGTGCTAGACAAATATCATGCACTTGTGAAATAATATATCTTGG 1800
QY 1801 CCTTTGGAATGAAGGCTAATCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
DB 1801 CCTTTGGAATGAAGGCTAATCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
QY 1861 TTAACACAAAGCCCGGAGACCTTCTTACGCGCTTCAAGAGAGAGCAAGAAAGAGAG 1920
DB 1861 TTAACACAAAGCCCGGAGACCTTCTTACGCGCTTCAAGAGAGAGCAAGAAAGAGAG 1920
QY 1921 GGTCTACTTTCATTTGGGTGAAAGAGATCAAGTGGCAAGACCCAGATCAAGTCTGTAG 1980
DB 1921 GGTCTACTTTCATTTGGGTGAAAGAGATCAAGTGGCAAGACCCAGATCAAGTCTGTAG 1980
QY 1981 AGCCATACACCAAGCAGCAGCTGACAAACATGTTCATTTGCTGAAATATATATGCGCTATA 2040
DB 1981 AGCCATACACCAAGCAGCAGCTGACAAACATGTTCATTTGCTGAAATATATATGCGCTATA 2040

[illegible]

```

STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,091
FILING DATE: 06-MAY-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/369,796
FILING DATE: 06-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2869 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
IMMEDIATE SOURCE:
LIBRARY: splenic/thymic
CLONE: Murine 19sf6
FEATURE:
NAME/KEY: CDS
LOCATION: 69..2378
US-08-852-091-11

```

Db 361 CAATGGAATTTGCCGATCTGTGCCCGATGCCGTGTGGGAAGATCTCGCTCTCCAGA 420
Oy 421 CGGCACCCAGCGAGCCCAAGAGGGGCCAGGCCAACCCACAGCCGCCGTAGTGA 480
Db 421 CGGCACCCAGCGAGCCCAAGAGGGGCCAGGCCAACCCACAGCCGCCGTAGTGA 480
Oy 481 CAGAGAAGCAGATGTTGGAGCAGCATCTTGAAGATGTCGGGAAGCGAGTGCAGATC 540
Db 481 CAGAGAAGCAGATGTTGGAGCAGCATCTTGAAGATGTCGGGAAGCGAGTGCAGATC 540
Oy 541 TAGAAGAGAAATGAAGTGTGTGGAGAACCTCCAGAGCAGCATTTGATTTTCACCTACAAAA 600
Db 541 TAGAAGAGAAATGAAGTGTGTGGAGAACCTCCAGAGCAGCATTTGATTTTCACCTACAAAA 600
Oy 601 CCTCAAGAGCCCAAGAGACATGCAGATGTGAATGGAACAACAGCTCTGTAGCCAGAC 660
Db 601 CCTCAAGAGCCCAAGAGACATGCAGATGTGAATGGAACAACAGCTCTGTAGCCAGAC 660
Oy 661 AGAAGATGCAGCAGCTGGAACAGATGCTACAGCCCTGAGACAGATGCGGAGAAGCATTTG 720
Db 661 AGAAGATGCAGCAGCTGGAACAGATGCTACAGCCCTGAGACAGATGCGGAGAAGCATTTG 720
Oy 721 TGAATGAGCTGGGGGGCTCTTGTACAGCAATGAGATACGTGCAGAAACACTGATGATG 780
Db 721 TGAATGAGCTGGGGGGCTCTTGTACAGCAATGAGATACGTGCAGAAACACTGATGATG 780
Oy 781 AAGAGCTGGCTGACTGGAAGAGCGCGCAGAGATCGCTGCATTCGAGGGCCCTCCACACA 840
Db 781 AAGAGCTGGCTGACTGGAAGAGCGCGCAGAGATCGCTGCATTCGAGGGCCCTCCACACA 840
Oy 841 TGTGCTGAGCCCTGTGGAAGAACTGATTAATCTTATGAGAGATTCACATTCAGACCC 900
Db 841 TGTGCTGAGCCCTGTGGAAGAACTGATTAATCTTATGAGAGATTCACATTCAGACCC 900
Oy 901 GCCAACAAATTAAGAACTGAGAGAGCTGCAGCAGAAAGTGTCTACAGAGCGGACCTTA 960
Db 901 GCCAACAAATTAAGAACTGAGAGAGCTGCAGCAGAAAGTGTCTACAGAGCGGACCTTA 960
Oy 961 TCGTGCGAGCAGCGGCGCATCTGGAGAGAGATCGTGGAGCTTTGAGAAGCTTAATGA 1020
Db 961 TCGTGCGAGCAGCGGCGCATCTGGAGAGAGATCGTGGAGCTTTGAGAAGCTTAATGA 1020
Oy 1021 AGATGCTCTGTGTGTGAGAGCGGAGCCCTGCATGCCATGCACCCGAGCGGCCCTTAG 1080
Db 1021 AGATGCTCTGTGTGTGAGAGCGGAGCCCTGCATGCCATGCACCCGAGCGGCCCTTAG 1080
Oy 1081 TCATCAAGACTGCTGTCCAGTTTACACAGAAAGTCAAGTTGCTGTCTGAAGT 1140
Db 1081 TCATCAAGACTGCTGTCCAGTTTACACAGAAAGTCAAGTTGCTGTCTGAAGT 1140
Oy 1141 TGAATTTACAGCTTAAATTTAAAGTGTGATTTGATTAAGACTCTGGGGATGTTCTGCC 1200
Db 1141 TGAATTTACAGCTTAAATTTAAAGTGTGATTTGATTAAGACTCTGGGGATGTTCTGCC 1200
Oy 1201 TCAGAGGGTCTCGGAAATTTAACTTCTGGGACAGAAACAAAAAGTATGAATGAGG 1260
Db 1201 TCAGAGGGTCTCGGAAATTTAACTTCTGGGACAGAAACAAAAAGTATGAATGAGG 1260
Oy 1261 AGTTAAACAAGCGAGCTGTCTGACAGATTCAAGCAGCTGACCTTAGAGAGAGAT 1320
Db 1261 AGTTAAACAAGCGAGCTGTCTGACAGATTCAAGCAGCTGACCTTAGAGAGAGAT 1320
Oy 1261 AGTTAAACAAGCGAGCTGTCTGACAGATTCAAGCAGCTGACCTTAGAGAGAGAT 1320
Db 1261 AGTTAAACAAGCGAGCTGTCTGACAGATTCAAGCAGCTGACCTTAGAGAGAGAT 1320
Oy 1321 GTGGAGATGAGGCGCGGCAATTTGATGCTCTCTTATGCTGATGAGAGAGCTGCACC 1380
Db 1321 GTGGAGATGAGGCGCGGCAATTTGATGCTCTCTTATGCTGATGAGAGAGCTGCACC 1380
Oy 1381 TGATCACCCTTCGAGACTGAGGTGTACCAAGGCGCTCAAGATTGACCTAGAGACCACT 1440
Db 1381 TGATCACCCTTCGAGACTGAGGTGTACCAAGGCGCTCAAGATTGACCTAGAGACCACT 1440
Oy 1441 CCTTGGCAGTTGTGTATCTCCAAACATCTGTGCAGATGCCAAATGCTTGGGATCAATCC 1500
Db 1441 CCTTGGCAGTTGTGTATCTCCAAACATCTGTGCAGATGCCAAATGCTTGGGATCAATCC 1500

Oy 1501 TGTGTATAACATGCTGACCAATTAACCCCAAGAAAGTGAACCTTCTCACTAACCCCAA 1560
Db 1501 TGTGTATAACATGCTGACCAATTAACCCCAAGAAAGTGAACCTTCTCACTAACCCCAA 1560
Oy 1561 TTGGAACCTTGGAGCAAGTGGCCGAGGTGCTCAGTGGAGTTCTGTCCAGCAGAAC 1620
Db 1561 TTGGAACCTTGGAGCAAGTGGCCGAGGTGCTCAGTGGAGTTCTGTCCAGCAGAAC 1620
Oy 1621 GAGGCTGACCATTCGAGCAGCTGACAAACGCTGGCTGAGAAAGTCTTAAGGCGCTGTGA 1680
Db 1621 GAGGCTGACCATTCGAGCAGCTGACAAACGCTGGCTGAGAAAGTCTTAAGGCGCTGTGA 1680
Oy 1681 ACTACTAGGCTGTCAGATGACATGGGCTAAATTTCTCAAGAAACATGGCTGCAAGG 1740
Db 1681 ACTACTAGGCTGTCAGATGACATGGGCTAAATTTCTCAAGAAACATGGCTGCAAGG 1740
Oy 1741 GCTTCTCTCTGTGGTCTGAGCAATATCATGACCTTGTGAAAGATATATCTTG 1800
Db 1741 GCTTCTCTCTGTGGTCTGAGCAATATCATGACCTTGTGAAAGATATATCTTG 1800
Oy 1801 CCTTTGGAATGAAGGTACATCATGCTTTCATCAGCAAGAGCCGAGCGGCATCC 1860
Db 1801 CCTTTGGAATGAAGGTACATCATGCTTTCATCAGCAAGAGCCGAGCGGCATCC 1860
Oy 1861 TAAGCAAAAGCCCGGAGACCTTCTACTGCGCTTCAGCGAAGCAGCAAAAGAGAG 1920
Db 1861 TAAGCAAAAGCCCGGAGACCTTCTACTGCGCTTCAGCGAAGCAGCAAAAGAGAG 1920
Oy 1921 GGGTCACTTCTACTTGGGTGAAAGAGACATCACTGSCAAGACCAGATCCAGTCTTAG 1980
Db 1921 GGGTCACTTCTACTTGGGTGAAAGAGACATCACTGSCAAGACCAGATCCAGTCTTAG 1980
Oy 1981 AGCCATTAACCAAGCAGCAGCTGACAAACATGCTATTTGCTGAATATCATGGGCTATA 2040
Db 1981 AGCCATTAACCAAGCAGCAGCTGACAAACATGCTATTTGCTGAATATCATGGGCTATA 2040
Oy 2041 AGATCATGATGCGACCAACATCTGCTGTCTCCACTTGTCTACTACCCGACATTC 2100
Db 2041 AGATCATGATGCGACCAACATCTGCTGTCTCCACTTGTCTACTACCCGACATTC 2100
Oy 2101 CCAAGAGAGAGCAATTTGAAAGTACTGTAGGCCCGGAGCCGAGAGCAGGACCCCAAGCCG 2160
Db 2101 CCAAGAGAGAGCAATTTGAAAGTACTGTAGGCCCGGAGAGCAGGACCCCAAGCCG 2160
Oy 2161 ACCGAGTAGTGTGCCCCGCTACTGAAAGACCAAGTTCATCTGTGTACACCAAGCACT 2220
Db 2161 ACCGAGTAGTGTGCCCCGCTACTGAAAGACCAAGTTCATCTGTGTACACCAAGCACT 2220
Oy 2221 GCAGCAATACCATGACCTGCGGATGTCGCCCGGACCTTATGATTCATGATGAGCTTTG 2280
Db 2221 GCAGCAATACCATGACCTGCGGATGTCGCCCGGACCTTATGATTCATGATGAGCTTTG 2280
Oy 2281 GAAATTAACGCTGAAGGTGCTGAGCCCTCAGCAGAGAGGAGTTGAGTCCGTACGTTTG 2340
Db 2281 GAAATTAACGCTGAAGGTGCTGAGCCCTCAGCAGAGAGGAGTTGAGTCCGTACGTTTG 2340
Oy 2341 ACATGATCTGACCTCGGAGTGTCTACCTCCCATGTGAGAGGCTGAAACCAAGAGCT 2400
Db 2341 ACATGATCTGACCTCGGAGTGTCTACCTCCCATGTGAGAGGCTGAAACCAAGAGCT 2400
Oy 2401 GCAGAGAGCTGACTTGAAGACACCTGCCCCGCTGCACACCCCTAAGCAGGCAACCCATA 2460
Db 2401 GCAGAGAGCTGACTTGAAGACACCTGCCCCGCTGCACACCCCTAAGCAGGCAACCCATA 2460
Oy 2461 TCGTCTGAACCTCTAACTTTGTGTCTCAGATTTTTTTTTTAATTTCTACTTGTCT 2520
Db 2461 TCGTCTGAACCTCTAACTTTGTGTCTCAGATTTTTTTTTTAATTTCTACTTGTCT 2520
Oy 2521 ATCTTTGGCAATCTGGCACTTTTAAAGAGAAATAGTAGTGGGTGATTAAC 2580
Db 2521 ATCTTTGGCAATCTGGCACTTTTAAAGAGAAATAGTAGTGGGTGATTAAC 2580

Db 721 TGAATGAGCTGGCGGGCTCTTGTACCAATGAGATGCTGACAGACACTGACTGATG 780
Oy 781 AAGAGCTGGCTGACTGGAAGAGCGCGCAGAGATCGCTGATGAGAGCCCTCCCAACA 840
Db 781 AAGAGCTGGCTGACTGGAAGAGCGCGCAGAGATCGCTGATGAGAGCCCTCCCAACA 840
Oy 841 TCTGCTGGAGCCGTGTGAAAACCTGATTAATTCATTAGCAATCTCAACTTCAGACCC 900
Db 841 TCTGCTGGAGCCGTGTGAAAACCTGATTAATTCATTAGCAATCTCAACTTCAGACCC 900
Oy 901 GCCAACAATTAAGAACTGGAGAGCTGCAGCGCAAGAAAGTGTCCCTACAGAGGCCACCTTA 960
Db 901 GCCAACAATTAAGAACTGGAGAGCTGCAGCGCAAGAAAGTGTCCCTACAGAGGCCACCTTA 960
Oy 961 TCGTGCAGACACCGCCCATGCTGGAGAGAGAGATCGTGGAGCTGTTCAAACTTAATGA 1020
Db 961 TCGTGCAGACACCGCCCATGCTGGAGAGAGATCGTGGAGCTGTTCAAACTTAATGA 1020
Oy 1021 AGAGTGCCTTCGTGGTGGAGCGGCGAGCCCTGCATGCCCCATGCCGAGCCGCGCTTAG 1080
Db 1021 AGAGTGCCTTCGTGGTGGAGCGGCGAGCCCTGCATGCCCCATGCCGAGCCGCGCTTAG 1080
Oy 1081 TCATCAAGACGTGCTGCTGCTTACCAAGTCAAGTGTGCTGCTGCTCAATTTCTGAGT 1140
Db 1081 TCATCAAGACGTGCTGCTGCTTACCAAGTCAAGTGTGCTGCTGCTCAATTTCTGAGT 1140
Oy 1141 TGAATTTATCAGCTTAAATTAAGTGTGCTTATTAAGTCTGAGAGCTTGGGGATTTGCTGCC 1200
Db 1141 TGAATTTATCAGCTTAAATTAAGTGTGCTTATTAAGTCTGAGAGCTTGGGGATTTGCTGCC 1200
Oy 1201 TCAGAGGCTCTCGGAATTTAACATTTCTGGCGAGACACAAAGTATGAACATGAGG 1260
Db 1201 TCAGAGGCTCTCGGAATTTAACATTTCTGGCGAGACACAAAGTATGAACATGAGG 1260
Oy 1261 AGTAAACAAGCGAGCGCTGCTGCAGAGTTCAAGCACCTGACCCCTTGGGAGCGAGAT 1320
Db 1261 AGTAAACAAGCGAGCGCTGCTGCAGAGTTCAAGCACCTGACCCCTTGGGAGCGAGAT 1320
Oy 1321 GTGGGAATGAGAGCGCTGCGCAATTTGTGATGCTCTTGTATGCTGATGAGAGCTGCACC 1380
Db 1321 GTGGGAATGAGAGCGCTGCGCAATTTGTGATGCTCTTGTATGCTGATGAGAGCTGCACC 1380
Oy 1381 TGATCACCTTCGAGATGAGGTGTACCAAGCGCTCAAGTGTGAGTACCTAGAGCCCACT 1440
Db 1381 TGATCACCTTCGAGATGAGGTGTACCAAGCGCTCAAGTGTGAGTACCTAGAGCCCACT 1440
Oy 1441 CCTTGCAGTTGTGATGCTCAACAATCTGTGCAGATGCCAAATGCTTGGGCATCAATCC 1500
Db 1441 CCTTGCAGTTGTGATGCTCAACAATCTGTGCAGATGCCAAATGCTTGGGCATCAATCC 1500
Oy 1501 TGTGCTATTAACATGCTGACCAATTAACCCCAAGACGTGAACCTTCTTCACTAAGCCGCAA 1560
Db 1501 TGTGCTATTAACATGCTGACCAATTAACCCCAAGACGTGAACCTTCTTCACTAAGCCGCAA 1560
Oy 1561 TTGGAACCTGGAGCAAGTGGCGGAGGTGCTCAGCTGGAGTTCCTGCGCAACCAAGC 1620
Db 1561 TTGGAACCTGGAGCAAGTGGCGGAGGTGCTCAGCTGGAGTTCCTGCGCAACCAAGC 1620
Oy 1621 GAGGGCTGAGCATGAGCAGCTGACAAAGCTGCTGAGAAAGCTCCTAGGGCTGATGGA 1680
Db 1621 GAGGGCTGAGCATGAGCAGCTGACAAAGCTGCTGAGAAAGCTCCTAGGGCTGATGGA 1680
Oy 1681 ACTAGCTAGGGTGTCAATGATCAATGGGCTAATCTCTGAAGAAACATGCTGGCAAGG 1740
Db 1681 ACTAGCTAGGGTGTCAATGATCAATGGGCTAATCTCTGAAGAAACATGCTGGCAAGG 1740
Oy 1741 GCTTCTCTCTCTGGTCTGGCTAGACAATATCATGACCTTGTAAAAAATATATCTTGG 1800
Db 1741 GCTTCTCTCTCTGGTCTGGCTAGACAATATCATGACCTTGTAAAAAATATATCTTGG 1800
Oy 1801 CCTTTTGAATGAAGGTTACATCATGGGTTTCATCAGCAAGAGCGGAGCGGCCATCC 1860
Db 1801 CCTTTTGAATGAAGGTTACATCATGGGTTTCATCAGCAAGAGCGGAGCGGCCATCC 1860

Oy 1861 TAAGCACAAGCCCCGGGCACTTCTCTACTGCGCTTCAGCGAGAGCAGCAAGAAGAG 1920
Db 1861 TAAGCACAAGCCCCGGGCACTTCTCTACTGCGCTTCAGCGAGAGCAGCAAGAAGAG 1920
Oy 1921 GGGTCACTTTTCACTTGGTGGAAAAAGACATCAGTGGCAAGACCCAGATCCAGTGTAG 1980
Db 1921 GGGTCACTTTTCACTTGGTGGAAAAAGACATCAGTGGCAAGACCCAGATCCAGTGTAG 1980
Oy 1981 AGCCATACACCAGCAGAGCTGAACAACATGCTCAATTTGCTGAATTCATCATGGCTTGA 2040
Db 1981 AGCCATACACCAGCAGAGCTGAACAACATGCTCAATTTGCTGAATTCATCATGGCTTGA 2040
Oy 2041 AGATCATGATGATGCGACCAACATCTGCTGTGCTGCACTTGTCTACCTTAACCCGACATTC 2100
Db 2041 AGATCATGATGATGCGACCAACATCTGCTGTGCTGCACTTGTCTACCTTAACCCGACATTC 2100
Oy 2101 CCAGGAGAGGCAATTTGGAAGTACTGTAGGCCCGAGAGCCAGAGACCCCGAAGCG 2160
Db 2101 CCAGGAGAGGCAATTTGGAAGTACTGTAGGCCCGAGAGCCAGAGACCCCGAAGCG 2160
Oy 2161 ACCCAGTAGTGTGCTGCGCGGATGCTGAAGACCAATTCATGCTGTGACACCAAGACCT 2220
Db 2161 ACCCAGTAGTGTGCTGCGCGGATGCTGAAGACCAATTCATGCTGTGACACCAAGACCT 2220
Oy 2221 GCAGCAATACCAATGACCTGCGGATGCTGCGCGGATGCTGCGGATGCTGCGGATGCTGCG 2280
Db 2221 GCAGCAATACCAATGACCTGCGGATGCTGCGCGGATGCTGCGGATGCTGCGGATGCTGCG 2280
Oy 2281 GAATTAAGCGTGAAGGTGCTGACGCTTACGACGAGAGGCGAGTTGATGCTGCTGAGCTTG 2340
Db 2281 GAATTAAGCGTGAAGGTGCTGACGCTTACGACGAGAGGCGAGTTGATGCTGCTGAGCTTG 2340
Oy 2341 ACATGATCTGCTGCTGCGAGTGTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Db 2341 ACATGATCTGCTGCTGCGAGTGTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
Oy 2401 GCAGAGAGTGAATTTGAGACACCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 2401 GCAGAGAGTGAATTTGAGACACCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Oy 2461 TCGTCTGAACCTCTTCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Db 2461 TCGTCTGAACCTCTTCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
Oy 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAAATGAGTGAAGTGGGTGATTAAC 2580
Db 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAAATGAGTGAAGTGGGTGATTAAC 2580
Oy 2581 TGTATGTAAGAGAGAGACCTCTGAGTCTGGGATGGGCTGAGAGCAGAGGAGGC 2640
Db 2581 TGTATGTAAGAGAGAGACCTCTGAGTCTGGGATGGGCTGAGAGCAGAGGAGGC 2640
Oy 2641 AAAGGGAACACCTCTGCTGCGCGGCTGCGCGGCTGCTTTCACACACTGGGGGTTGG 2700
Db 2641 AAAGGGAACACCTCTGCTGCGCGGCTGCGCGGCTGCTTTCACACACTGGGGGTTGG 2700
Oy 2701 TTGTTAGACAAGTGCCTCTGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Db 2701 TTGTTAGACAAGTGCCTCTGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
Oy 2761 ACCCATCTCTGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Db 2761 ACCCATCTCTGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
Oy 2821 TAGGACTAAGCCAGAGAGTCTCTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2869
Db 2821 TAGGACTAAGCCAGAGAGTCTCTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2869

RESULT 4
US-08-956-652-11
: Sequence 11, Application US/08956652

Patent No. 6013475
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,652
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,185
FILING DATE: 11-MAR-1994
APPLICATION NUMBER: US 07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2869 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
IMMEDIATE SOURCE:
LIBRARY: splenic/lymfc
CLONE: Murine 19s16
FEATURE:
NAME/KEY: CDS
LOCATION: 69..2378
US-08-956-652-11

Query Match 100.0%; Score 2869; DB 3; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GCCGCGACAGCCAGCGCGCGCTCAGCCCGAGACAGTCGAGACCCCTGACT 60

|||||
Db 1 GCCGCGACAGCCAGCGCGCGCTCAGCCCGAGACAGTCGAGACCCCTGACT 60
OY 61 GCAGCAGATGCGCTCACTGGAACCAAGCTGCAGCAGCTTGACACACGCTACTGAAGCAGC 120
|||||
Db 61 GCAGCAGATGCGCTCACTGGAACCAAGCTGCAGCAGCTTGACACACGCTACTGAAGCAGC 120
OY 121 TGCACAGCTGTACAGGACACGTTCCCATGAGCTGCGGAGCTTCTTGACCTTGGA 180
|||||
Db 121 TGCACAGCTGTACAGGACACGTTCCCATGAGCTGCGGAGCTTCTTGACCTTGGA 180
OY 181 TTGAGAGTCAAGACTGGGCAATATGACCCAGCAAGAGTACATGCCAGCTTGTGTTTC 240
|||||
Db 181 TTGAGAGTCAAGACTGGGCAATATGACCCAGCAAGAGTACATGCCAGCTTGTGTTTC 240
OY 241 ATATCTCTTGGGTAATTGACCAATATACCCGATTCCTCCAGAGTCCAAATGTC 300
|||||
Db 241 ATATCTCTTGGGTAATTGACCAATATACCCGATTCCTCCAGAGTCCAAATGTC 300
OY 301 TCTATCAGCAACACTTTCGAAGATCAAGCAGTTTCTGCAGCAGGTATCTTGAGAAC 360
|||||
Db 301 TCTATCAGCAACACTTTCGAAGATCAAGCAGTTTCTGCAGCAGGTATCTTGAGAAC 360
OY 361 CAATGCAATTTGCCGATGCTGCGCCGATGCTCTGTGGAGAGTCTGCTCTCTAGA 420
|||||
Db 361 CAATGCAATTTGCCGATGCTGCGCCGATGCTCTGTGGAGAGTCTGCTCTCTAGA 420
OY 421 CGCGAGCAGCGCAGCCAGCAAGGGGGCCAGCCCAACCAACCCAGCCGCTACTGCA 480
|||||
Db 421 CGCGAGCAGCGCAGCCAGCAAGGGGGCCAGCCCAACCAACCCAGCCGCTACTGCA 480
OY 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTCAGATGTCGCGAGCAGTGCAGATC 540
|||||
Db 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTCAGATGTCGCGAGCAGTGCAGATC 540
OY 541 TAGAAGCAAAATGAGAGTGTGAGAACCTCCAGACACTTTGATTCACTACAAAA 600
|||||
Db 541 TAGAAGCAAAATGAGAGTGTGAGAACCTCCAGACACTTTGATTCACTACAAAA 600
OY 601 CCTCAAGACCCCAAGGACATGACAGATCTGAATGGAACCAACCACTGTGACCCAGAC 660
|||||
Db 601 CCTCAAGACCCCAAGGACATGACAGATCTGAATGGAACCAACCACTGTGACCCAGAC 660
OY 661 AGAAGATGACAGCAGCTGGAACAGATGCTCACAGCCCTTGACAGCAGATCGGAGAAATG 720
|||||
Db 661 AGAAGATGACAGCAGCTGGAACAGATGCTCACAGCCCTTGACAGCAGATCGGAGAAATG 720
OY 721 TGAAGTACGCTGGGGGGCTCTTGTACAGCAATGAGTACGTCAGAAACACTGACTGATG 780
|||||
Db 721 TGAAGTACGCTGGGGGGCTCTTGTACAGCAATGAGTACGTCAGAAACACTGACTGATG 780
OY 781 AAGAGTGCCTGACTGGAAGAGCGGGCCAGAGATCGGTCATGGGAGCCCTGCCACA 840
|||||
Db 781 AAGAGTGCCTGACTGGAAGAGCGGGCCAGAGATCGGTCATGGGAGCCCTGCCACA 840
OY 841 TCTGCTGACCGCTCTGGAACACTGATACTTATGAGAGATCTCAACTTCAGACCC 900
|||||
Db 841 TCTGCTGACCGCTCTGGAACACTGATACTTATGAGAGATCTCAACTTCAGACCC 900
OY 901 GCCAACAAATTAAGAACTGAGAGAGCTGCAGAGAAAGTCTCTACAGGGGACCCCTA 960
|||||
Db 901 GCCAACAAATTAAGAACTGAGAGAGCTGCAGAGAAAGTCTCTACAGGGGACCCCTA 960
OY 961 TCGTGCAGCAGCGCCCATGCTGAGAGAGAGATCGTGGAGCTGTTTCAGAACTTAATGA 1020
|||||
Db 961 TCGTGCAGCAGCGCCCATGCTGAGAGAGAGATCGTGGAGCTGTTTCAGAACTTAATGA 1020
OY 1021 AGAGTCCCTTCTGATGAGAGCGGCGCAGCCCTGATGCGCATGCAACCGGAGCCGCTTAG 1080
|||||
Db 1021 AGAGTCCCTTCTGATGAGAGCGGCGCAGCCCTGATGCGCATGCAACCGGAGCCGCTTAG 1080
OY 1081 TCATCAAGAGCTGTGTCAGATTACCAAGCAAGTCAAGTTGCTGTGTAATTTCTGAGCT 1140
|||||

Db 1081 TCATCAAGACTGCTGTCAGCTTTACACGAAGCTCAGCTTCTGCTCAAAATTTCTGAGT 1140
Qy 1141 TGAATTAATCAAGCTTTAAATTAAGTGTGATGATTAAGAGCTGTGGGATGTGCTGCCC 1200
Db 1141 TGAATTAATCAAGCTTTAAATTAAGTGTGATGATTAAGAGCTGTGGGATGTGCTGCCC 1200
Qy 1201 TCGAGGCTCTCGGAATTTAACTTCTGGGACAGAACACAAAAGTATGAACATGAGAGG 1260
Db 1201 TCGAGGCTCTCGGAATTTAACTTCTGGGACAGAACACAAAAGTATGAACATGAGAGG 1260
Qy 1261 AGTCTAACACAGCGAGCCCTGCTCTGACAGATTCAAGACCTGACCTTAGAGGAGAGAT 1320
Db 1261 AGTCTAACACAGCGAGCCCTGCTCTGACAGATTCAAGACCTGACCTTAGAGGAGAGAT 1320
Qy 1321 GTGGGAATGAGAGCCGCTGCTGATGCTCTCTGATGCTGATGAGAGAGCTGCAACC 1380
Db 1321 GTGGGAATGAGAGCCGCTGCTGATGCTCTCTGATGCTGATGAGAGAGCTGCAACC 1380
Qy 1381 TGATCACCCTTCGACACTGAGCTGTACCAAGCCCAAGGCTCAAGATTGACCTAGAGCCACT 1440
Db 1381 TGATCACCCTTCGACACTGAGCTGTACCAAGCCCAAGGCTCAAGATTGACCTAGAGCCACT 1440
Qy 1441 CCTTGCCAGTTGGTGATCTCCAAATCTGTGATGCTGCAATGCTTGGGCTCAATCC 1500
Db 1441 CCTTGCCAGTTGGTGATCTCCAAATCTGTGATGCTGCAATGCTTGGGCTCAATCC 1500
Qy 1501 TGTGGTATTAACATGCTGACCAATTAACCCCAAGAGTGAATCTTCTCACTAGAGCCCAA 1560
Db 1501 TGTGGTATTAACATGCTGACCAATTAACCCCAAGAGTGAATCTTCTCACTAGAGCCCAA 1560
Qy 1561 TTGGAACTTGGGACCAAGTGGCCGAGGTGCTCAGCTGGCAGATTCCTGCTCACCAAGC 1620
Db 1561 TTGGAACTTGGGACCAAGTGGCCGAGGTGCTCAGCTGGCAGATTCCTGCTCACCAAGC 1620
Qy 1621 GAGGGCTGAGCATGAGACAGCTGCAACGCTGGCTGAGAGCTCTTAGGGCTGTGTGA 1680
Db 1621 GAGGGCTGAGCATGAGACAGCTGCAACGCTGGCTGAGAGCTCTTAGGGCTGTGTGA 1680
Qy 1681 ACTACTAGGGTGTCAATCATGATGAGGCTTAATCTGCAAGAAACATGGCTGGCAAG 1740
Db 1681 ACTACTAGGGTGTCAATCATGATGAGGCTTAATCTGCAAGAAACATGGCTGGCAAG 1740
Qy 1741 GCTTCTCTCTCTGGGCTGTGCTAGACAAATATCATGCACTTGTGAAAAAGTATATCTTGG 1800
Db 1741 GCTTCTCTCTCTGGGCTGTGCTAGACAAATATCATGCACTTGTGAAAAAGTATATCTTGG 1800
Qy 1801 CCTTTGGAATGAAGGGTACATCATGGGTTTCATGACAAAGAGCGGAGTGGCCATCC 1860
Db 1801 CCTTTGGAATGAAGGGTACATCATGGGTTTCATGACAAAGAGCGGAGTGGCCATCC 1860
Qy 1861 TAAACCAAAAGCCCCCGGACCTTCTTACTGCGCTTCAGGAGAGAGCAAAAGAGAG 1920
Db 1861 TAAACCAAAAGCCCCCGGACCTTCTTACTGCGCTTCAGGAGAGAGCAAAAGAGAG 1920
Qy 1921 GGGTCACTTTCTACTGGGTGAGAAAGACATCACTGCGAGAGACCAATCCAGTCTGAG 1980
Db 1921 GGGTCACTTTCTACTGGGTGAGAAAGACATCACTGCGAGAGACCAATCCAGTCTGAG 1980
Qy 1981 AGCCATACACCAAGCAGCAGCTGAAACAATGATTTGCGAANAATCATATGGGCTATA 2040
Db 1981 AGCCATACACCAAGCAGCAGCTGAAACAATGATTTGCGAANAATCATATGGGCTATA 2040
Qy 2041 AGATCATGATGCGAGCAACATCTGTGTCTCCACTTGTCTTACCTTACCCGACATTC 2100
Db 2041 AGATCATGATGCGAGCAACATCTGTGTCTCCACTTGTCTTACCTTACCCGACATTC 2100
Qy 2101 CCAAGGAGAGGCAATTTGGAAGTACTGTAGCCCGAGAGCAGAGACACCCCAAGCGG 2160
Db 2101 CCAAGGAGAGGCAATTTGGAAGTACTGTAGCCCGAGAGCAGAGACACCCCAAGCGG 2160
Qy 2161 ACCGAGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2161 ACCGAGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220

Qy 2221 GCACCAATACATTTGACCTGCCCAGATGTCGCCGACCTTTAGATTTCATGATGAGTTTG 2280
Db 2221 GCACCAATACATTTGACCTGCCCAGATGTCGCCGACCTTTAGATTTCATGATGAGTTTG 2280
Qy 2281 GAAATTAACGGTGAAGGCTGCTGAGCCCTGACGAGAGAGGACATTTGAGTCCGCTACGTTTG 2340
Db 2281 GAAATTAACGGTGAAGGCTGCTGAGCCCTGACGAGAGAGGACATTTGAGTCCGCTACGTTTG 2340
Qy 2341 ACATGATCTGACCTCGAGTGTGCTACCTCCCTCATGAGAGCTGAAACAGAACT 2400
Db 2341 ACATGATCTGACCTCGAGTGTGCTACCTCCCTCATGAGAGCTGAAACAGAACT 2400
Qy 2401 GCAGAGAGCTGCTGAGACACCTGCCCCGCTGCTCCACCCCTAAGCAGCCGACATTA 2460
Db 2401 GCAGAGAGCTGCTGAGACACCTGCCCCGCTGCTCCACCCCTAAGCAGCCGACATTA 2460
Qy 2461 TCGTCTGAACCTCTTAACCTTTGTGTTTCAGATTTTCTTCTTCTTCTTCTTCTTCT 2520
Db 2461 TCGTCTGAACCTCTTAACCTTTGTGTTTCAGATTTTCTTCTTCTTCTTCTTCTTCT 2520
Qy 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGGGTATTAAC 2580
Db 2521 ATCTTTGGGCAATCTGGGCACTTTTAAAGAGAGAAATGAGTGTGGGTATTAAC 2580
Qy 2581 TGTATGTAAAGAGAGAGACCTGAGTGTGGGATGGGCTGAGAGCAGAGAGAGGAGC 2640
Db 2581 TGTATGTAAAGAGAGAGACCTGAGTGTGGGATGGGCTGAGAGCAGAGAGAGGAGC 2640
Qy 2641 AAAGGGAAACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Db 2641 AAAGGGAAACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
Qy 2701 TTGTTAACAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 2760
Db 2701 TTGTTAACAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 2760
Qy 2761 ACCCATTTCTGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 2820
Db 2761 ACCCATTTCTGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 2820
Qy 2821 TAGGACTTAAGCCCGAGAGGTTCTCTTTTAAATTAATTAATTAATTAATTAATTAATTAAT 2869
Db 2821 TAGGACTTAAGCCCGAGAGGTTCTCTTTTAAATTAATTAATTAATTAATTAATTAATTAAT 2869

RESULT 5
US-08-956-869-11
; Sequence 11, Application us/08956869
; Patent No. 6030808
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956, 869

Oy	1501	TGCGTATATACATGCTGCACCATTAACCCCAAGCTGACCTTCTCACTAAAGCCGCAA	1560
Db	1501	TGTGTATATACATGCTGCACCAATTAACCCCAAGACCTGAACCTTCTCACTAAAGCCGCAA	1560
Oy	1561	TTGGAAACCTGGGAGACCAGAGTGGGCGAGAGTCTCAGCTGGGAGGTTCTGTCCACACCAAGC	1620
Db	1561	TTGGAAACCTGGGAGACCAGAGTGGGCGAGAGTGGGCGAGAGTGGGAGGTTCTGTCCACACCAAGC	1620
Oy	1621	GAGGCGTGAACATCCGAGCAGCTGACAAAGCTGGCTGAGAAAGCTCCTAAGGCGCTGGGTGA	1680
Db	1621	GAGGCGTGAACATCCGAGCAGCTGACAAAGCTGGCTGAGAAAGCTCCTAAGGCGCTGGGTGA	1680
Oy	1681	ACTACTCAGGGGTGTCAGATCACATGGGCTAAATTTCTGCAAGAAAACATGGCTGGCAAG	1740
Db	1681	ACTACTCAGGGGTGTCAGATCACATGGGCTAAATTTCTGCAAGAAAACATGGCTGGCAAG	1740
Oy	1741	GCTTCTCCTTCTGGGTCTGGCTGAGCAATATCATCGACCTGTGGAAAAAGTATATCTGG	1800
Db	1741	GCTTCTCCTTCTGGGTCTGGCTGAGCAATATCATCGACCTGTGGAAAAAGTATATCTGG	1800
Oy	1801	CCCTTGGATGAAGGATACATCATGGGTTTCATCAGCAAGAGCGGAGCGGCGCATCC	1860
Db	1801	CCCTTGGATGAAGGATACATCATGGGTTTCATCAGCAAGAGCGGAGCGGCGCATCC	1860
Oy	1861	TAAACACAAGCCCCCGGGACCTTCTACTGCGCTTCAAGAGAGCAGCAAGAAGAGAG	1920
Db	1861	TAAAGCAAAAGCCCCCGGGACCTTCTACTGCGCTTCAAGAGAGCAGCAAGAAGAGAG	1920
Oy	1921	GGGTCACTTTCACTTGGGTGGAAGAAAGACATCAGTGGCAAGACCACATCAGTCTAG	1980
Db	1921	GGGTCACTTTCACTTGGGTGGAAGAAAGACATCAGTGGCAAGACCACATCAGTCTAG	1980
Oy	1981	AGCCATACACCAAGCAGCAGCTGAACACATGTCATTTGCTGAATCATCATGGGCTATA	2040
Db	1981	AGCCATACACCAAGCAGCAGCTGAACACATGTCATTTGCTGAATCATCATGGGCTATA	2040
Oy	2041	AGATCATGATGGGACCAACATCCTGGTGTCTCAGTTGCTACCTTACCCTCCGACATTC	2100
Db	2041	AGATCATGATGGGACCAACATCCTGGTGTCTCAGTTGCTACCTTACCCTCCGACATTC	2100
Oy	2101	CCAAGGAGGAGGCAATTTGGAAGTACTGTAGGCCCGAGAGCCAGAGCACCCCGAAGCCG	2160
Db	2101	CCAAGGAGGAGGCAATTTGGAAGTACTGTAGGCCCGAGAGCCAGAGCACCCCGAAGCCG	2160
Oy	2161	ACCCAGTATGCTGCCCTTACTCTGAAGACCAAGTTCATCTGTGTACACCAAGCAAGCT	2220
Db	2161	ACCCAGTATGCTGCCCTTACTCTGAAGACCAAGTTCATCTGTGTACACCAAGCAAGCT	2220
Oy	2221	GCAGCATATACATTTGACCTCCGATGTGCCCCCGACCTTTAGATTCATTTGACGATTTG	2280
Db	2221	GCAGCATATACATTTGACCTCCGATGTGCCCCCGACCTTTAGATTCATTTGACGATTTG	2280
Oy	2281	GAATATACGCTGAAGGTGTCTGAGGCCCTCAGCAGAGAGGCAATTTGAGTCCGTCACGTTTG	2340
Db	2281	GAATATACGCTGAAGGTGTCTGAGGCCCTCAGCAGAGAGGCAATTTGAGTCCGTCACGTTTG	2340
Oy	2341	ACATGATCTGACCTCGGAGTGTCTACTCTCCCATGTGAGAGAGCTGAACACGAAGCT	2400
Db	2341	ACATGATCTGACCTCGGAGTGTCTACTCTCCCATGTGAGAGAGCTGAACACGAAGCT	2400
Oy	2401	GCAGAGACGTAATGAGACACCTGCCCGGTGTCCACACCCTTAAGCAGGCCAGCCCATTA	2460
Db	2401	GCAGAGACGTAATGAGACACCTGCCCGGTGTCCACACCCTTAAGCAGGCCAGCCCATTA	2460
Oy	2461	TGCGTGAACCTCCTACTCTGTGGTGTCCAAATTTTTTTTTTAATTTCCATCTTGCT	2520
Db	2461	TGCGTGAACCTCCTACTCTGTGGTGTCCAAATTTTTTTTTTTTTTAATTTCCATCTTGCT	2520
Oy	2521	ATCTTTGGCAATCTGGGCACTTTTAAAAAGAAATGATGATGTTGGGTATATAAC	2580
Db	2521	ATCTTTGGCAATCTGGGCACTTTTAAAAAGAAATGATGATGTTGGGTATATAAC	2580
Oy	2581	TGTTATCTAAAGCAGAGACCTTGAGTCTGGGATGTGGCTGAGACAGAGCGAGAGC	2640

[illegible]

RESULT 6
 US-08-948-547-11
 Sequence 11, Application US/08948547
 Patent No. 6124118
 GENERAL INFORMATION:
 APPLICANT: Darnell Jr., James E.
 APPLICANT: Schindler, Christian W.
 APPLICANT: Fu, Xian-Yuan
 APPLICANT: Wen, Xilong
 APPLICANT: Zhong, Zhong
 TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
 TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Klauber & Jackson
 STREET: 111 Hackensack Avenue
 CITY: " Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/948,547
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/212,185
 FILING DATE: 11-MAR-1994
 APPLICATION NUMBER: US/07/980,498
 FILING DATE: 23-NOV-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/854,296
 FILING DATE: 19-MAR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO US93/02569
 FILING DATE: 19-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/126,588
 FILING DATE: 24-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-073 CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201 487-5800
 TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION FOR SFO ID NO: 11:
 SEQUENCE CHARACTERISTICS:

LENGTH: 2869 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: mouse
IMMEDIATE SOURCE:
LIBRARY: splenic/thymic
CLONE: Murine 19sf6
FEATURE:
NAME/KEY: CDS
LOCATION: 69..2378
US-08-948-547-11

Query Match 100.0%; Score 2869; DB 3; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCCGACACCGACGCCGCCAGTGGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT 60
DB 1 GCCCGACACCGACGCCGCCAGTGGGCTCAGCCCGGAGACAGTGCAGACCCCTGACT 60
OY 61 GCACAGAGATGGCTCAGTGGAAACAGCTGCAGAGCTGGACACACGGTACTGGAAGCAGC 120
DB 61 GCACAGAGATGGCTCAGTGGAAACAGCTGCAGAGCTGGACACACGGTACTGGAAGCAGC 120
OY 121 TGCACAGCTGTAGACGCACAGCTTCCCATGAGAGCTGGGCACTTCTGGACCTTGA 180
DB 121 TGCACAGCTGTAGACGCACAGCTTCCCATGAGAGCTGGGCACTTCTGGACCTTGA 180
OY 181 TTGAGAGTCAGACTGGGCGATGATGAGCCAGCAAGATGCATGACCACTTGGTTTC 240
DB 181 TTGAGAGTCAGACTGGGCGATGATGAGCCAGCAAGATGCATGACCACTTGGTTTC 240
OY 241 ATATATCTCTGGTGAATTTGACCAGCAATATAGCCGATTCCTGCAAGAGTCCCATATGCC 300
DB 241 ATATATCTCTGGTGAATTTGACCAGCAATATAGCCGATTCCTGCAAGAGTCCCATATGCC 300
OY 301 TCTATCAGACACACCTTGAAGAAATCAAGCAATTTCTGCACAGACAGTATCTTGAAGAAC 360
DB 301 TCTATCAGACACACCTTGAAGAAATCAAGCAATTTCTGCACAGACAGTATCTTGAAGAAC 360
OY 361 CAATGGAATTTGCCCGATCGTGGCCCGATGCTGTGGAGAGATCTTGCTCTCCAGA 420
DB 361 CAATGGAATTTGCCCGATCGTGGCCCGATGCTGTGGAGAGATCTTGCTCTCCAGA 420
OY 421 CGGACAGCCAGGACGCCAGCAAGGGGGCCAGGCCAACCCCAACAGCCCGCTAGTGA 480
DB 421 CGGACAGCCAGGACGCCAGCAAGGGGGCCAGGCCAACCCCAACAGCCCGCTAGTGA 480
OY 481 CAGAGAACAGCAATGTTGGAGAGCATCTTCAGAGATGTCGGGAAGGAGATGGAGATC 540
DB 481 CAGAGAACAGCAATGTTGGAGAGCATCTTCAGAGATGTCGGGAAGGAGATGGAGATC 540
OY 541 TAGAAGCAAAATGAAGGTGTGGAGAACTTCAGAGAGCACTTTGATTCAACTACAAAA 600
DB 541 TAGAAGCAAAATGAAGGTGTGGAGAACTTCAGAGAGCACTTTGATTCAACTACAAAA 600
OY 601 CCCCTCAAGAGCAAGAGACATGTCAGATCTGAAATGAAAAACCAAGCTCTTGACCCAGAC 660
DB 601 CCCCTCAAGAGCAAGAGACATGTCAGATCTGAAATGAAAAACCAAGCTCTTGACCCAGAC 660
OY 661 AGAAGATGCACAGCTGGAACAGATGCTCACAGCCTTGAGCAGATGGGAGAGCATTTG 720
DB 661 AGAAGATGCACAGCTGGAACAGATGCTCACAGCCTTGAGCAGATGGGAGAGCATTTG 720
OY 721 TGAAGTACTGGCGGGCTTTGTGAGCAATGAGATGCTGCAGAGACACTGACTGATG 780
DB 721 TGAAGTACTGGCGGGCTTTGTGAGCAATGAGATGCTGCAGAGACACTGACTGATG 780

OY 781 AAGAGCTGGCTGACTGGAAGAGGGGCCAGAGATCAGCTGATGGAGGCTTCCCAACA 840
DB 781 AAGAGCTGGCTGACTGGAAGAGGGGCCAGAGATCAGCTGATGGAGGCTTCCCAACA 840
OY 841 TCTGCTGTGACCGCTCTGGAAACATGATTAATCTTATAGCAGAAATCTCAACTTCAGACCC 900
DB 841 TCTGCTGTGACCGCTCTGGAAACATGATTAATCTTATAGCAGAAATCTCAACTTCAGACCC 900
OY 901 GCCAACAATTAAGAAACCTGAGAGAGCTGCACAGAAAGTGTCTACAAAGCGGACCTTA 960
DB 901 GCCAACAATTAAGAAACCTGAGAGAGCTGCACAGAAAGTGTCTACAAAGCGGACCTTA 960
OY 961 TCGTGCAGACCGGCCCATGCTGAGAGAGAGATCGTGAAGCTTTGAGAACTTAATGA 1020
DB 961 TCGTGCAGACCGGCCCATGCTGAGAGAGAGATCGTGAAGCTTTGAGAACTTAATGA 1020
OY 1021 AGAGTGCCTTCTGCTGAGAGAGCGGACCGCTGCATGCCATGCAACCGGACCGCTTAG 1080
DB 1021 AGAGTGCCTTCTGCTGAGAGAGCGGACCGCTGCATGCCATGCAACCGGACCGCTTAG 1080
OY 1081 TCATCAGAGACTGCTGTCCAGTTTACACAGAAAGTCAAGTCTGTGCAAAATTTCTGAGT 1140
DB 1081 TCATCAGAGACTGCTGTCCAGTTTACACAGAAAGTCAAGTCTGTGCAAAATTTCTGAGT 1140
OY 1141 TGAATTTACAGCTTAAATTTAAAGTGTGATTAAGACTGTGGGATGTTGCTGCC 1200
DB 1141 TGAATTTACAGCTTAAATTTAAAGTGTGATTAAGACTGTGGGATGTTGCTGCC 1200
OY 1201 TCAGAGGCTCTCGGAATTTTAACATTTTGAGGACAGAACACAAAAGTATACATGAGAG 1260
DB 1201 TCAGAGGCTCTCGGAATTTTAACATTTTGAGGACAGAACACAAAAGTATACATGAGAG 1260
OY 1261 AGTCTAACAGGCGAGCGCTGCTCAGAGTTCAAGCAACCTGACCTTAGGAGACAGAT 1320
DB 1261 AGTCTAACAGGCGAGCGCTGCTCAGAGTTCAAGCAACCTGACCTTAGGAGACAGAT 1320
OY 1321 GTGGAAATGAGGCGGTCGCAATTTGTATGCTCTTGTATGCTGAGAGCTGCACC 1380
DB 1321 GTGGAAATGAGGCGGTCGCAATTTGTATGCTCTTGTATGCTGAGAGCTGCACC 1380
OY 1381 TGATCACCCTTCAGAGCTGAGGTTGACCAAGGCGCTCAAGATGACCTTAGACCCACT 1440
DB 1381 TGATCACCCTTCAGAGCTGAGGTTGACCAAGGCGCTCAAGATGACCTTAGACCCACT 1440
OY 1441 CCTTGGCAGTTGTGATGTCACAACTGTGTGATGATGCCAAATGCTTGGGCATCAATCC 1500
DB 1441 CCTTGGCAGTTGTGATGTCACAACTGTGTGATGATGCCAAATGCTTGGGCATCAATCC 1500
OY 1501 TGTGTTAATACATGCTGACCAATTAACCCCAAGAACGTGAATCTTCACTAAGCCGCAA 1560
DB 1501 TGTGTTAATACATGCTGACCAATTAACCCCAAGAACGTGAATCTTCACTAAGCCGCAA 1560
OY 1561 TTGGAACCTGGAGCAAGTGGCCGAGGTGCTCAGCTGGAGATTTCTGTCACCAACCAAG 1620
DB 1561 TTGGAACCTGGAGCAAGTGGCCGAGGTGCTCAGCTGGAGATTTCTGTCACCAACCAAG 1620
OY 1621 GAGGGCTGAGCATGACAGCTGACCAACGCTGAGAGAGCTCTGAGGCGCTGGGTGA 1680
DB 1621 GAGGGCTGAGCATGACAGCTGACCAACGCTGAGAGAGCTCTGAGGCGCTGGGTGA 1680
OY 1681 ACTACTGAGGCTGCATCATCATGAGGCTAAATCTGCAAGAAACATGAGCTGGCAAG 1740
DB 1681 ACTACTGAGGCTGCATCATCATGAGGCTAAATCTGCAAGAAACATGAGCTGGCAAG 1740
OY 1741 GCTTCTCTTCTGGGCTGTGGCTAGACAAATATCATGCACTTTGTAAGAAATATATCTTG 1800
DB 1741 GCTTCTCTTCTGGGCTGTGGCTAGACAAATATCATGCACTTTGTAAGAAATATATCTTG 1800
OY 1801 CCTTTTGAATGAAGGTTACTCATGSGTTTCATCAGCAAGAGAGCGGAGCGCATCC 1860
DB 1801 CCTTTTGAATGAAGGTTACTCATGSGTTTCATCAGCAAGAGAGCGGAGCGCATCC 1860

1861 TAAGCAAAAGCCCCGGGACCTTCTACTGCGCTTCAAGCGAGACGCAAGAGAG 1920
1861 TAAGCAAAAGCCCCGGGACCTTCTACTGCGCTTCAAGCGAGACGCAAGAGAGAG 1920
1921 GGGTCACCTTTCACCTGGTGGTGAAGAGACATCAGTGGCAAGACCCAGATCCACTGTAG 1980
1921 GGGTCACCTTTCACCTGGTGGTGAAGAGACATCAGTGGCAAGACCCAGATCCACTGTAG 1980
1981 AGCCATACCAAGACAGACACCTGCAACATGTCATTTTCTGAAATATCATATGGGCTATA 2040
1981 AGCCATACCAAGACAGACACCTGCAACATGTCATTTTCTGAAATATCATATGGGCTATA 2040
2041 AGATCATGATGGAGACCAATATCTGCTGCTCCTACTGCTACTGCTACTGCTGCTGCTATTC 2100
2041 AGATCATGATGGAGACCAATATCTGCTGCTCCTACTGCTACTGCTACTGCTGCTGCTATTC 2100
2101 CCAAGAGAGAGGCAATTTGAAAGTACTGTAGCCCGAGACCCAGAGACCCCGAAGCCG 2160
2101 CCAAGAGAGAGGCAATTTGAAAGTACTGTAGCCCGAGACCCAGAGACCCCGAAGCCG 2160
2161 ACCAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
2161 ACCAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
2221 GCAGCAATACCATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
2221 GCAGCAATACCATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
2281 GAAATTAACGGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
2281 GAAATTAACGGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
2341 ACATGATCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2341 ACATGATCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2401 GCAGAGAGCTGATTTGAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
2401 GCAGAGAGCTGATTTGAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
2461 TCGCTGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2461 TCGCTGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2521 ATCTTTGGGCAATCTGCGACATTTTAAAGAGAGAAATGAGTGTGGTGAATAAC 2580
2521 ATCTTTGGGCAATCTGCGACATTTTAAAGAGAGAAATGAGTGTGGTGAATAAC 2580
2581 TGTATGTAAGAGAGAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
2581 TGTATGTAAGAGAGAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
2641 AAAGGGGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
2641 AAAGGGGAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
2701 TTGTTAGACAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
2701 TTGTTAGACAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
2761 ACCCATTTGGAAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
2761 ACCCATTTGGAAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
2821 TAGGACTAAGCCAGAGAGGTTCTCTTTAAATTAATAAAAAAAAAAAAAAAAAAAAAA 2869
2821 TAGGACTAAGCCAGAGAGGTTCTCTTTAAATTAATAAAAAAAAAAAAAAAAAAAAAA 2869

RESULT 7
PCT-US95-17025-11
; Sequence 11, Application PC/TUS9517025
; GENERAL INFORMATION:

APPLICANT: James E. Darnell, Jr.
APPLICANT: Zilong Wen
APPLICANT: Curt M. Horvath
APPLICANT: Zhong Zhong
TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17025
FILING DATE: 28-DEC-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/369,796
FILING DATE: 06-JAN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SFO ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2869 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse
IMMEDIATE SOURCE:
LIBRARY: splenic/lymnic
CLONE: Murine 19sf6
FEATURE:
NAME/KEY: CDS
LOCATION: 69..2378
PCT-US95-17025-11

Query Match 100.0%; Score 2869; DB 5; Length 2869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCCCGACACGACGCGCGGCGAGTGGGCTCAGCGCCGAGACAGTGGAGACCCCTACT 60
1 GCCCGACACGACGCGCGGCGAGTGGGCTCAGCGCCGAGACAGTGGAGACCCCTACT 60
61 GCAGGAGATGGCTGAGTGAACGAGTGGAGAGCTGGACACAGCTACTGGAAGCAGC 120
61 GCAGGAGATGGCTGAGTGAACGAGTGGAGAGCTGGACACAGCTACTGGAAGCAGC 120
121 TGCACACAGCTGTACAGCAGACGTTCCCATGAGAGCTGGCAGTTCCTGGACCTTGA 180
121 TGCACACAGCTGTACAGCAGACGTTCCCATGAGAGCTGGCAGTTCCTGGACCTTGA 180
181 TTGAGAGTCAAGACTGGGCGATATGAGCCAGCAAAAGATCAATGCAACGTTGGTGTTC 240
181 TTGAGAGTCAAGACTGGGCGATATGAGCCAGCAAAAGATCAATGCAACGTTGGTGTTC 240

Db 181 TTGAGAGTCAGAGCTGGGCATATGACAGCAAGAGTCACATGCCACGTTGGTTC 240
OY 241 ATATATCTTTGGGTGAATTTGACACCAATATATAGCCATTCCTGCAAGAGTCCATGTC 300
Db 241 ATATATCTTTGGGTGAATTTGACACCAATATATAGCCATTCCTGCAAGAGTCCATGTC 300
OY 301 TCTATGAGCACAACCTTCGAAGATCAAGCAGTTTCTGAGAGCAGGTATCTTGAGAGC 360
Db 301 TCTATGAGCACAACCTTCGAAGATCAAGCAGTTTCTGAGAGCAGGTATCTTGAGAGC 360
OY 361 CATGGAATATGCGCGGATGTGGCCGATGCTGTGGGAGAGTCTGCGCTCTCCAGA 420
Db 361 CATGGAATATGCGCGGATGTGGCCGATGCTGTGGGAGAGTCTGCGCTCTCCAGA 420
OY 421 CGGACGCCAGCGGACCCAGCAAGGGGGCCAGGCCCAACCCAGCCAGCCGCTAGTGA 480
Db 421 CGGACGCCAGCGGACCCAGCAAGGGGGCCAGGCCCAACCCAGCCAGCCGCTAGTGA 480
OY 481 CAGAGAAAGCAGATGTTGGAGCAGCATCTTCAAGATGTCGGGAAGCGAGTGCAGATC 540
Db 481 CAGAGAAAGCAGATGTTGGAGCAGCATCTTCAAGATGTCGGGAAGCGAGTGCAGATC 540
OY 541 TAGAACGAAATGAAGTGTGTGAGAACCTCCAGAGCAGCTTTGATTTCACTACAAAA 600
Db 541 TAGAACGAAATGAAGTGTGTGAGAACCTCCAGAGCAGCTTTGATTTCACTACAAAA 600
OY 601 CCTCAAGAGCCAGAGACATGACAGATCTGAATGAAACCAACAGTCTGTGACACAGAC 660
Db 601 CCTCAAGAGCCAGAGACATGACAGATCTGAATGAAACCAACAGTCTGTGACACAGAC 660
OY 661 AGAAGATGACAGCTGGAACAGATGCTACAGCCCTGAGCCAGAGTCCGAGAGCATTG 720
Db 661 AGAAGATGACAGCTGGAACAGATGCTACAGCCCTGAGCCAGAGTCCGAGAGCATTG 720
OY 721 TGAATGAGCTGGCGGGGCTCTTCTCAGCAATGAGTACGTGAGAAAGACATGATG 780
Db 721 TGAATGAGCTGGCGGGGCTCTTCTCAGCAATGAGTACGTGAGAAAGACATGATG 780
OY 781 AAGAGCTGGCTGCTGGAAGAGCGGCCAGAGATCGGTGTCATGCGAGCGCTCCCAACA 840
Db 781 AAGAGCTGGCTGCTGGAAGAGCGGCCAGAGATCGGTGTCATGCGAGCGCTCCCAACA 840
OY 841 TGTGCTGACCGCTGGAAGAACTGATTAATTGAGAGATCTCAACTTCAGACC 900
Db 841 TGTGCTGACCGCTGGAAGAACTGATTAATTGAGAGATCTCAACTTCAGACC 900
OY 901 GCCAACAAATTAGAAAGCTGAGAGCTGAGAGAAAGTGTCTCAAGAGGAGCCCTA 960
Db 901 GCCAACAAATTAGAAAGCTGAGAGCTGAGAGAAAGTGTCTCAAGAGGAGCCCTA 960
OY 961 TCGTGACGACCGCGCCCATGCTGAGAGAGAGATCGTGAAGCTGTTCAAGAACTTAATGA 1020
Db 961 TCGTGACGACCGCGCCCATGCTGAGAGAGAGATCGTGAAGCTGTTCAAGAACTTAATGA 1020
OY 1021 AGAGTCCCTTCTGTGTGTGAGAGCGGACCCCTGCAATGCCATGCAACCGGACCCCTTAG 1080
Db 1021 AGAGTCCCTTCTGTGTGTGAGAGCGGACCCCTGCAATGCCATGCAACCGGACCCCTTAG 1080
OY 1081 TCAATCAAGAGCTGTGTGACATTTACCAAGAAAGTCAAGTGTGTGCAAAATTTCTGAGT 1140
Db 1081 TCAATCAAGAGCTGTGTGACATTTACCAAGAAAGTCAAGTGTGTGCAAAATTTCTGAGT 1140
OY 1141 TGAATATACAGCTTTAAATTTAAAGTGTGATTTAAAGATCTGTGGGATGTGTGCCCC 1200
Db 1141 TGAATATACAGCTTTAAATTTAAAGTGTGATTTAAAGATCTGTGGGATGTGTGCCCC 1200
OY 1201 TCGAGAGGCTCTGGAAATTTAATTTCTGGGACGAAACAAAAAGTATGAATGAGG 1260
Db 1201 TCGAGAGGCTCTGGAAATTTAATTTCTGGGACGAAACAAAAAGTATGAATGAGG 1260
OY 1261 AGCTTAACAACGAGCAGCTGTCTGCAAGAGTTCAAGCAACCTGACCTTTAGAGCAGAGAT 1320
Db 1261 AGCTTAACAACGAGCAGCTGTCTGCAAGAGTTCAAGCAACCTGACCTTTAGAGCAGAGAT 1320

OY 1321 GTGGAAATGAGAGCGCGTCCCAATTTGTGATCCCTCTTGATCGTGACTGAGAGCTGCACC 1380
Db 1321 GTGGAAATGAGAGCGCGTCCCAATTTGTGATCCCTCTTGATCGTGACTGAGAGCTGCACC 1380
OY 1381 TGATCACCTTTCAGAGCTGAGAGTGTACCAAGAGCCCAAGATTGAGACTAGAGACCCACT 1440
Db 1381 TGATCACCTTTCAGAGCTGAGAGTGTACCAAGAGCCCAAGATTGAGACTAGAGACCCACT 1440
OY 1441 CCTTGCCAGTTGTGTGATCTCCCAACATCTGTGATGCCCAATGCTTGGGCAATATCC 1500
Db 1441 CCTTGCCAGTTGTGTGATCTCCCAACATCTGTGATGCCCAATGCTTGGGCAATATCC 1500
OY 1501 TGTGTATTAACATGTGTGACCAATTAACCCAGAAAGCTGAATCTTCACTAAGCCGCAA 1560
Db 1501 TGTGTATTAACATGTGTGACCAATTAACCCAGAAAGCTGAATCTTCACTAAGCCGCAA 1560
OY 1561 TTGGAACCTGGAGCCAAAGTGGCGAGGTGCTCAGCTGGAGTTCTGTCCAGCCAGAGC 1620
Db 1561 TTGGAACCTGGAGCCAAAGTGGCGAGGTGCTCAGCTGGAGTTCTGTCCAGCCAGAGC 1620
OY 1621 GAGGCTGAGCATGAGACAGCTGACAAACGCTGCTGAGAAAGCTCTAGGCTGTGTGTA 1680
Db 1621 GAGGCTGAGCATGAGACAGCTGACAAACGCTGCTGAGAAAGCTCTAGGCTGTGTGTA 1680
OY 1681 ACTACTAGAGGTGTCAAGATGACATGGGCTAAATTTGCAAGAAAGAAATGATGGCTGGCAGG 1740
Db 1681 ACTACTAGAGGTGTCAAGATGACATGGGCTAAATTTGCAAGAAAGAAATGATGGCTGGCAGG 1740
OY 1741 GCTTCTCTCTCTGGGTCTGCTGACAAATATGCACTTCGCTTGAAGAAATATATCTTG 1800
Db 1741 GCTTCTCTCTCTGGGTCTGCTGACAAATATGCACTTCGCTTGAAGAAATATATCTTG 1800
OY 1801 CCTTTTGAATGAAGGCTACATGAGTTTCATCAGCAAGAGAGCGGAGCCGACATCC 1860
Db 1801 CCTTTTGAATGAAGGCTACATGAGTTTCATCAGCAAGAGAGCGGAGCCGACATCC 1860
OY 1861 TAAACAAAGCCCGCGGACCTTCTACTGCGCTTACAGAGAGAGCAAGAAAGAGAG 1920
Db 1861 TAAACAAAGCCCGCGGACCTTCTACTGCGCTTACAGAGAGAGCAAGAAAGAGAG 1920
OY 1921 GGGTCACTTTCACCTGGGTGAAAAAGACATCAGTGGAGACCCAGATCCAGTGTAG 1980
Db 1921 GGGTCACTTTCACCTGGGTGAAAAAGACATCAGTGGAGACCCAGATCCAGTGTAG 1980
OY 1981 AGCCATACCAAGCAGCAGCTGAAACATGTCATTTGCTGAATTCATATGAGGCTATA 2040
Db 1981 AGCCATACCAAGCAGCAGCTGAAACATGTCATTTGCTGAATTCATATGAGGCTATA 2040
OY 2041 AGATCATGATGAGGACCAACATGCTGTCTCAGCTTGTCTACCTTACCCGACATTC 2100
Db 2041 AGATCATGATGAGGACCAACATGCTGTCTCAGCTTGTCTACCTTACCCGACATTC 2100
OY 2101 CCAAGAGAGAGCATTTTGAAAGTACTGTAGGCCGAGAGCCAGAGACCCCGAAGCCG 2160
Db 2101 CCAAGAGAGAGCATTTTGAAAGTACTGTAGGCCGAGAGCCAGAGACCCCGAAGCCG 2160
OY 2161 ACCCAGTAGTGTGCGCCCTTACCTGGAAGACCAAGTTCACTGTGTGACACCAAGCAGCT 2220
Db 2161 ACCCAGTAGTGTGCGCCCTTACCTGGAAGACCAAGTTCACTGTGTGACACCAAGCAGCT 2220
OY 2221 GCAGCAATACCATTTGACCTCCGATGTCGCCCGCACTTTAGATTCATTTGATGACAGTTTG 2280
Db 2221 GCAGCAATACCATTTGACCTCCGATGTCGCCCGCACTTTAGATTCATTTGATGACAGTTTG 2280
OY 2281 GAAATTAAGGTGAAGGTGCTGAGCCCTCAGCAGAGAGGCGATTGATGCTGACAGTTTG 2340
Db 2281 GAAATTAAGGTGAAGGTGCTGAGCCCTCAGCAGAGAGGCGATTGATGCTGACAGTTTG 2340
OY 2341 ACATGATCTGAGCTCGGAGTGTGCTAACCTGCCCAATGTGAGAGAGCTGAACAGAGCT 2400
Db 2341 ACATGATCTGAGCTCGGAGTGTGCTAACCTGCCCAATGTGAGAGAGCTGAACAGAGCT 2400

QY 2401 GCAGAGCGTGAATGAGACACCTGCCCGCTGCTCCACCCTTAAGACGCCAACCCCATATA 2460
| | | | |
DB 2401 GCAGAGCGTGAATGAGACACCTGCCCGCTGCTCCACCCTTAAGACGCCAACCCCATATA 2460
QY 2461 TCGCTGAAAGCTCTTAACCTTTGCTGCTCCAGATTTTTTTTTTAATTTCTACTTCTGCT 2520
| | | | |
DB 2461 TCGCTGAAAGCTCTTAACCTTTGCTGCTCCAGATTTTTTTTTTAATTTCTACTTCTGCT 2520
QY 2521 ATCTTGGGCAATCTGGGCACTTTTAAAGAGAAATAGTAGTGAGTGAGTAAAC 2580
| | | | |
DB 2521 ATCTTGGGCAATCTGGGCACTTTTAAAGAGAAATAGTAGTGAGTGAGTAAAC 2580
QY 2581 TGTATGTAAGAGAGAGACCTCTGAGTGGGATGGGCTGAGAGCAGAGAGAGC 2640
| | | | |
DB 2581 TGTATGTAAGAGAGAGACCTCTGAGTGGGATGGGCTGAGAGCAGAGAGAGC 2640
QY 2641 AAGGGGAAACCTCTGCTGCTGCCCGCTGCTCCCTCTTTTTCAGCAGCTCGGGGCTTG 2700
| | | | |
DB 2641 AAGGGGAAACCTCTGCTGCTGCCCGCTGCTCCCTCTTTTTCAGCAGCTCGGGGCTTG 2700
QY 2701 TTGTTAGACAAGTGCCTGCTGCTGCCCATGCTACCTGTTGCCCATCTGTGAGCTGAT 2760
| | | | |
DB 2701 TTGTTAGACAAGTGCCTGCTGCTGCCCATGCTACCTGTTGCCCATCTGTGAGCTGAT 2760
QY 2761 ACCCCATTTCTGGGAACCTCTGCTGCTGCTACCTTCAACCTTGAATATTCACATAGAAC 2820
| | | | |
DB 2761 ACCCCATTTCTGGGAACCTCTGCTGCTGCTACCTTCAACCTTGAATATTCACATAGAAC 2820
QY 2821 TAGACTTAAGCCAGAGGTTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTA 2869
| | | | |
DB 2821 TAGACTTAAGCCAGAGGTTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTA 2869

RESULT 8
US-09-288-461-82
; Sequence 82, Application US/09288461
; Patent No. 6139694
; GENERAL INFORMATION:
; APPLICANT: Karraas, James G.
; TITLE OF INVENTION: Antisense Oligonucleotide Modulation of STAT3
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0338
; CURRENT APPLICATION NUMBER: US/09/288,461
; CURRENT FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 82
; LENGTH: 2869
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (69)..(2381)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: U069222
; DATABASE ENTRY DATE: 1994-07-01
US-09-288-461-82

Query Match 99.8% Score 2864.2 DB 3 Length 2869:
Best Local Similarity 99.9% Pred. No. 0:
Matches 2866: Conservative 0; Mismatches 3; Indels 0; Gaps 0;

DB 121 tgcacagctgtacacgacacgllccccaatgagctgagcagllccttgcaccttga 180
QY 181 TTGAGAGTCAGAGCTGGCATATGACAGCCAGAAAGATCAGATGCCACCTGCTGTTTC 240
| | | | |
DB 181 ttgagagtcacagctgtgacatgacagccagcaaaagctacatgacagcttgcagcttgc 240
QY 241 ATATCTCTTGGGTGAATATTGACAGCAATATAGCCGATTTCTGCAAGATATCCAAATGTC 300
| | | | |
DB 241 ataactccttgggtgaatattgacagcaatatagcagatccttgcagagatccaaatgctc 300
QY 301 TGTATGACAGCAACCTTGGAAGATCAAGCAGTTTTCGAGAGAGATATCTTGAGAAC 360
| | | | |
DB 301 tctatgcagcaaaccttgcagaaatcaagcagatcttgcagagcagatccttgcagaaagc 360
QY 361 CAATGCAATTTGCCCGGATGCTGCCGATGCCGTTGGGAGAGACTGCTCCCTCTCCAGA 420
| | | | |
DB 361 caatgcaatttgcccgatgctgccgatgccgcttgggagagactgctccctctccaga 420
QY 421 CGGACGCGACGCGACGCCAGAGGGGGCCAGGCAACCCACAGCCGCGCTAGTGA 480
| | | | |
DB 421 cggacgcgacgcgacgccagagggggccagggcaacccacagccgcgctagtga 480
QY 481 CAGAGAGCAGAGATGTTGAGCAGCATCTTTCAGATGTCGGAAGCAGATGAGAGATC 540
| | | | |
DB 481 cagagagcagagatgttgagcagcatctttcagatgtcggaaagcagatgagagatc 540
QY 541 TAGAAGCAAGAAATGAAGGTGTGAGAACCTCCAGAGCAGCTTGTATTTCACTACAAA 600
| | | | |
DB 541 tagaagcaagaatgaaggtgtgagaacctccagagcagcttgtatttcaactacaaa 600
QY 601 CCTCAAGAGAGCAGAGAGATGAGATGATGATGGAACCAACCATCTGTGACAGAC 660
| | | | |
DB 601 cctcaagagagcagagagatgagatgatgatggaaaccaacctgtgacagac 660
QY 661 AGAAGATGACAGCAGCTGAGACAGATGCTCACAGCCTTGAGACAGATGCGAGAGACTTG 720
| | | | |
DB 661 agaagatgacagcagctgagacagatgctcacagccttgagacagatgcgagagacttg 720
QY 721 TGAGTGAAGCTGCGCGGCTCTTGTACCAATGAGATGAGTACGTCAGAGACACTGATG 780
| | | | |
DB 721 tgagtgaagctgcgcgctcttgtaccaatgagatgagtacgctcagagacactgatg 780
QY 781 AAGAGCTGCGTGAATGGAAGAGGGGCGCAGAGATCGGTGATGGAGGCGCTCCCAACA 840
| | | | |
DB 781 aagagctgctggaatggaaagagggcgccagagatcggtgatggagggcgctcccaaca 840
QY 841 TCTGCTGAGACCGTCTGGAAGAACTGATTAATCTTATAGCAGAAATCTCAACTTCAGACC 900
| | | | |
DB 841 tctgctgagaccgtctggaaagaaactgatataatcttataagcaaatctcaacttcagacc 900
QY 901 GCCAACAATTAAGAACTGAGAGAGCTGCAGCAGAAAGTCTCTACAGGGCAGCCTTA 960
| | | | |
DB 901 gccacaataatgaagaaactgagagagctgcagcagaaagtcttcaacaaggcgagacctta 960
QY 961 TCGTGCAGCAGCGGCCCATGCTGAGAGAGAGATCGTGGAGCTGTGCAAACTTAATGA 1020
| | | | |
DB 961 tcgtgcagcagcgcccatgctgagagagagatcgtggagctgtgcaaaacttaatga 1020
QY 1021 AGAGTGCCTTCTGCTGAGAGCGGCGACGCTTCATGCCCATGCCAGCGGCGCCTTAG 1080
| | | | |
DB 1021 agagtgccttctgctgagagcgggcgacgcttcattgcccatgccagcgcgcccttag 1080
QY 1081 TCAATCAGACTGTGTGCTCAGTTTACACGAAAGTCAAGTGTCTGTCTCAATTTCTCTAGT 1140
| | | | |
DB 1081 tcaatcagactgtgtgctcagtttaccgaaagtcaagtgctgtctcaatttctctagt 1140
QY 1141 TGAATTAACAGCTTAATAAGTGTGATTTGAATTAAGATCTGAGGATGTTGCTGCTGCC 1200
| | | | |
DB 1141 tgaattaacagcttaataagtgtgatTTGAATTAAGATCTGAGGATGTTGCTGCTGCC 1200
QY 1201 TCAGA TGTCTGCAAAATTTTACATTTCTGCGCACGAAACAAAGTGAATGAGATGAGG 1260
| | | | |
DB 1201 tcagatgtcttcgaaatlttaacatltctgagcgacgaacacaaagtgtgaaacatgagag 1260

OY	1261	AGCTAACACAGCGACGGCTGTGTCACAGAGTTCAAGCACTGACCTTAAAGGACAGAGAT	1320
Db	1261	agctcaaaacaagcgagacgtctgtccagagtlctaaagcaacctgaccccttagaggagcayagat	1320
OY	1321	GTGGGAATGGAGGCGCGGCAATTGGATGGCCCTTGATTCGTAGCTAGAGAGCGACAC	1380
Db	1321	gtgggaatggagagcgcgagccaattgtagtgctctcttgatctgtgactgagagctgcgacc	1380
OY	1381	TGATACCTTCGAGACTGAGGTGTACCCAAAGCCTCAAGATTGACCTAGAGACCCACT	1440
Db	1381	tgatacccttcgagactgaggtgtaccaccaagcgctcaagattgacctagagaccact	1440
OY	1441	CGTTGCCGAGTGTGTGATATCCACATGTCGTAGATGGCAATCGCTTGGGATCAATCC	1500
Db	1441	cgltgcagcgtctgtgtgatlctccaacatcgtccatgagatgcgcaatgctctggagatcaatcc	1500
OY	1501	TGTGTATACACTGCTGACCAATPACCCCAAGACGTGAACCTTCTCACTAAGCCGCCAA	1560
Db	1501	tgtgtataaactctgtaccaataaaccccaagaagctgaactcttcaactaagccgcraa	1560
OY	1561	TTGGAACTGGGAGACAGTAGTGCGAGGTGCTGAGCTGGAGTTCTGTGTCACAC	1620
Db	1561	ttggaaactgggagacagtagtgcgaggtgctgagctggagcttctgtccacacacgaac	1620
OY	1621	GAGGCTGACGACATGAGACAGCTGACAAAGCTGAGCTAGACAGCTCTGAGGCCTGGTGTGA	1680
Db	1621	gaggctgagacatcgagcagctgacaagcgtgctgagagctctcctlaaggccttgtytga	1680
OY	1681	ACTACTCAGGGTGTCAATACATGAGGCTTAAATTCTGCACAAACAAACATGCTGGCAAG	1740
Db	1681	actactcagggltcagatcacatctgtgtcaaatctctgcacaagaataactgctctgcaag	1740
OY	1741	GCTTCTCCTTGTGGGTGTGGCTAGACATATCTACACCTGTGAAAAAGTTATCTGG	1800
Db	1741	gcttctccttctgtgtctgtgctagacaatacatctgacctctgtgaaaaagatatcttgg	1800
OY	1801	CCCTTTGGAATGAAGGTTACATCATGCGTTTATCAGCAAGAGCGCGAGCGGCATCC	1860
Db	1801	ccctttggaatgaagggttacatcatggttctacacagaagagcgagcgccatcc	1860
OY	1861	TAAAGACAAACCCCCCGGACACCTCTACTCGCGTTACGCGAGAGCAAGCAAGAGAG	1920
Db	1861	taagacaacaagcccccgagcaecttctactctgagcttcagcgagagcaagaagaagag	1920
OY	1921	GGGTCACTTTCACTTGGGTGGAAAGGACATCAGTGGCAAGACCCAGATCCAGTTGTAG	1980
Db	1921	gggtcaacttctacttgggtgaaagaacatcagtgcgaagccagatccagctcgtag	1980
OY	1981	AGGCATACACCAACGACGACTGTAACAACTGTCAATTGCTGAAATCATATGGGCTATA	2040
Db	1981	agccatacaacaagcaagcagctgaaacaattgcatcttgcgaaatcatcatcagggctata	2040
OY	2041	AGATCATGAGTGCACACCAACATCTCGTGCTGTCAATTTGTACCTTACACCCGCATTC	2100
Db	2041	agatcatgagtgcgaccaaatactcgtgtctcactatgttctactcttaaccgcgacatc	2100
OY	2101	CCAGAGGAGGAGCACTTTGGAAAGTACTGTAGGCCGAGAGCCAGAGACACCCCGAAGCGC	2160
Db	2101	ccaaggaaggagcatcttgaaagtlactgtagccgagagccaagagacaccccgaaagcg	2160
OY	2161	ACCCAGTAGTGCAGCCCGTACCTGGAAGCAAGTCACTGTGTGACACCAACGACCT	2220
Db	2161	acccagtagtgcagcccgatcccgatccgaaagcaagltcatctgtgaaaccaagacct	2220
OY	2221	GCAGCAATACCATTTGACCTCGCATGTCCCGCCGCACTTTAATTCATTTGATGACGTTTG	2280
Db	2221	gcagcaatacatctgacctgcgagtgctcccccgcacatttagatcatgtgacgttg	2280
OY	2281	GAAATPACGGTGAAGGTGCTGAGCCTCTACAGACGAGAGGACGATTTTAGTGCCTCACTTTG	2340
Db	2281	gaaatpaaagtgaaagtgagctgagccctcagcagaggaagagtttagtgcgtccacglttg	2340

QY	2341	ACATGATCTGACCTCGAGAGTGCTGCTACTCCCCCATGTGAGAGACTGAAACCAACT	2400
DB	2341	acatgatictgaccctcgagatgltgctactccccaatgtagagatctgaaccataaact	2400
QY	2401	GCAGAGACTGACTTGAGAGACACTGGCCCGTCTCCACCCCTAAGCAGCGCAACCCATA	2460
DB	2401	gcagagacattgacttgagagacactggcccggtctccaccctaaagcagcgcaaccata	2460
QY	2461	TCGTCTGAAACCTCTPAACCTTTGTGTGTCAGATTTTTTTTTTAAATTTCTACTTCGCT	2520
DB	2461	tcgtctgaaccccttaacctgtgtgtccagatcttttttlaatttctcactctgct	2520
QY	2521	ATCTTTGGGCATCTGGGCACTTTTAAAGAGAGAAATGAGTGAGTGGGGTGAATTAAC	2580
DB	2521	atctttgggcaatctgggcaacttttaaaagagaaatgatlgaatgtagggtgaataac	2580
QY	2581	TGTTATGTAAAGAGAGAGACCTCTGAGTCTGGGATGGGGCTGACAGCAGAAGGAGGC	2640
DB	2581	tgttatgtaaaagagagacactctgagcttgagctgggctgacagcagaagagagc	2640
QY	2641	AAAGGGAGACCTCTCTGTCCTGCGCCGCTGCTCTTTTACAGAGCTGGGGGTTGG	2700
DB	2641	aaagggagacactctctgctcgcgcgcgcctcttctcaagaagctgggggttgg	2700
QY	2701	TGTTAGACAGACTGCTCTCTGTGTCGCCATGAGCTGTTGGCCCACTGTGAGCTGAT	2760
DB	2701	tgttagacagactgctctctgtgtgcccatgagctgttggcccaactgtgagctgat	2760
QY	2761	ACCCCATTTGTGGGAACCTCTGTGCTGACACTTTCACCTTGCTAATATCCACATAGAGC	2820
DB	2761	accccatcttggaacctctgtgctctgcaacttcaaccttgctaatatccatagaagc	2820
QY	2821	TGAGACTAGCCCGAGAGAGTCTCTTTAAATTAATTAATTAATTAATTAATTAATTAAT	2869
DB	2821	tagagactaagcccgagaggtctctcttctaataataaaaaaaaaaaaaaa	2869
RESULT	9		
US-09-364-970-9			
Sequence 9, Application US/09364970			
Patent No. 6235873			
GENERAL INFORMATION:			
APPLICANT: Bromberg, Jacqueline			
TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR			
TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING			
TITLE OF INVENTION: DYSPLASIA/PROLIFERATIVE CELLULAR CHANGES			
FILE REFERENCE: 600-1-252			
CURRENT APPLICATION NUMBER: US/09/364,970			
CURRENT FILING DATE: 1999-07-31			
NUMBER OF SEQ ID NOS: 10			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 9			
LENGTH: 2869			
TYPE: DNA			
ORGANISM: Mus musculus			
US-09-364-970-9			
Query Match	99.7%;	Score 2861;	DB 4; Length 2869;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 2864;	Conservative	0; Mismatches	5; Indels 0; Gaps 0;

QY 181 TTGAGAGTCAAGACTGGGATATGCGCCAGCAAGAGTGCATGCGACGTGGTTTC 240
|||||
Db 181 ttgagagtaagacatgggcatatgcagccaagatcaatgcacgttggtttc 240
QY 241 ATAAATCTTGGTGAAATTGACACCAATATAGCCGATTCTCGAAGAGTCCAAATGTCC 300
|||||
Db 241 ataatactcttgggtgaattgaccagcaatatagcgattctcgtcaagagttccaatgtcc 300
QY 301 TCTATCAGCACAACCTTCCAGAAATACAGACGTTTCTCGAGACAGGTATCTTGAGAGC 360
|||||
Db 301 tctatcagcacaaccccttcgaagaatacaagcagttctgcagagcaaglatcttgaagaagc 360
QY 361 CATGTGAAATTTGCCGATGCTGGCCGATGCTGTGGAGAGTCTGCCCTCTCCAGA 420
361 caatggaataatgcccggatgctggccgatgctgtggagagagctcgcgctctccaga 420
QY 421 CGGACGCCACGACGCCACCAAGGGGCCAGGCCAACCCACAGCCGCCGTAGTGA 480
|||||
Db 421 cggacgccacgagcccaagggggccaggccaaccaccaacagccgcgtagtgga 480
QY 481 CAGAGAGCAGCAGATGTTGAGCAGCATCTTACGATGTCCGAAAGCAGTGCAGGATC 540
|||||
Db 481 cagagagcagcagatgttggagcagcatcttcaggaatgtccggaaagcgagtgaggatc 540
QY 541 TAGAACGAAAATGAGGTGTGGAGAACCTCCAGACGACTTTGATTCACTACAAA 600
|||||
Db 541 tagaacgaaaatgaaagtgtggagaaacctcagagacgacttgaatcacaataaaa 600
QY 601 CCCTCAGAGCCAAAGGAGATGAGATGTAATGGAACCAACAGCTCTGTGACCGAGC 660
601 cccctcaagagccaagagaaatcgaagatctgaaatgaatacaacaaatctgtgcacgac 660
QY 661 AGAAGATGACAGCAGCTGGAACAGATGCTCACAGCCCTGGACCAAGTCGGAGAACTTG 720
|||||
Db 661 agaagatgacagcagctggaacagatgctcacagccctggaccaaagatcgagaaagcatltg 720
QY 721 TGAATGAGCTGGCGGGCTCTGTGCACAAATGAGTACGTACGTACGAAGACACTGACTGTG 780
|||||
Db 721 tgaatgagctggcgggctctgtgcacaaatgagtgatcgtaagaaagacactgtgactgtg 780
QY 781 AAGAGCTGGCTGACTGGAGAGGCGGCCAGAGATCGCTGCATCGGAGGCCCTCCCAACA 840
|||||
Db 781 aagagctggctgactggagagcgccgcagagatcgcgtgacatcggaagccctcccaaca 840
QY 841 TCTGCTGTGACCGCTGTGGAACACTGATTAATCTATTAGCAGATTCGAATTCAGACCC 900
|||||
Db 841 tctgctgtgaccgctgtggaataactgataacttcatatagcgaatctcaacttcagacc 900
QY 901 GCCAACAATTAAGAAATGAGAGAGCTGCAGCAAAAGTGTCTACAAGGCGACCTTA 960
901 gccacaataatlaagaatacttgaagagctgcagcagaagaatgtctcaacaaggcgaccccta 960
QY 961 TCGTGACAGCCGGCCCATGCTGTGAGAGAGAGATCGTGGAGCTGTTCAAACTTAATGA 1020
|||||
Db 961 tcgtgacagccggcccatgctgtgagagagatcgttgaagactgtltcagaataactlaaaga 1020
QY 1021 AGAGTGCTTGTGTGTGAGGCGGACCCCTGCATGCCCATGCACCCGGAGCGGCCCTTAG 1080
1021 agagtgcttgcgtgtgagggcgacccctgcagatgcacacccggagccggcccttaag 1080
QY 1081 TCATCAGACTGTGTCTCAGTTTACACGAAAGTTCAGTGTCTGTCTCAATTTCTGTAGT 1140
1081 tcatcagactgtgtctcagtttacacgaaagtctgctgtgttcaaatltctcgtagt 1140
QY 1141 TGAATTATCAGCTTAAATTTAAAGTGTGATTAAGACTCTGCGGATGTTGTGTGCC 1200
1141 tgaattatcagctttaaatttaaaagtgtgcatgataaagactctgggagatgtgtgtgcc 1200
QY 1201 TCAGAGGCTCTCGAAATTTAAATCTGTGGCAGCAACACAAAGTGTGAACATGAGAG 1260
1201 tcagaggtctctcgaaatttaaacatctcgtggcagcaacaacaaagtgtatgaacatggagg 1260

QY 1261 AGTCTAACACGGCAGCCTGTCTGACAGTTCAAGCAGCTTACGACCTTAGGAGCAGAGT 1320
1261 agtctaacaacgagccctgtctgcagagttcaagcacttgacccttaaggagcagagat 1320
QY 1321 GTGGGAATGAGCGCGGTGCCAATTTGTATGCCCTCTCTGTGATCTGATGTGAGGAGCTTCAC 1380
1321 gtgggaatgagcgcggtgcccaatttgtatgccctctctgtgattctgagatgagagctgcac 1380
QY 1381 TGATTCACCTTTCGAGACTGAGCTGTACCAACAAAGGCTTCAGATTTGACTTAGAGACCCACT 1440
1381 tgatcaccttctcgagactgagagtgtaaccacaagcctcaagaatttgacatagagaccact 1440
QY 1441 CCTTTTTCAGTTTGTGTGATCTCCAAACATCTGTTCAGATTCGCTTGAATTCGCTGATATCC 1500
1441 cctttttcagtttgtgtgattctccaaacatctgttcagatggccaatgcttggatcaatcc 1500
QY 1501 TGTGTATACATGCTTCACCAATTAACCCCAAGAACCTGAACTTTCTTCACTTAAGCCGCTCAA 1560
1501 tgtgtatacatgcttcaccaatlaaacccaagaagcgtgaactcttcaactlaagcgcgcaa 1560
QY 1561 TTGGAACCTTGGACCAAGTGCGCGAGTGCTCAGCTTGCGACTTCTGTCCACCAACCAAGC 1620
1561 ttggaaccttggaccaagtgcgagtgctcagctggcagttctgcaccaacaagc 1620
QY 1621 GAGGCTGAGCATGCGAGCAGCTGACAAAGCTGGCTGGAAGCTTCAGGCTGTGTGTA 1680
1621 gaggctgagcatgagcagctgacaaagctggctggaagcttcctlaaggccctgtgtga 1680
QY 1681 ACTACTAGGGGTGTACATCATCATGAGCTTAATTTCTGCAAGAAACATGCTGTGCAAGG 1740
1681 actactaggggtgtacatcatcatgagcttaatttctgcaagaataaataagctgtgcaag 1740
QY 1741 GCTTCTGCTTCTGGGCTGTGCTACACAAATATGATCGACTTGTGAAAGAAATATATCTTG 1800
1741 gcttctgcttctgggctgtgctacacaaatgatcgacttgtgaaagaaatataatcttg 1800
QY 1801 CCTTTTGAATGAAGGCTTACATCATGAGCTTTCATCAGCAAGAGCGGAGCGGCCATTC 1860
1801 ccttttgaatgaaggcttacctatgagcttctcatcagcaagagcgaggccctc 1860
QY 1861 TAAACCAAAACCCCCCGCGACCTTCTACTGCGCTTCACGCAAGACGCAAAAGAAAGAG 1920
1861 taacacaagaagcccccgacacttctcctacatgctgtcctcaagagagagcaagaagag 1920
QY 1921 GGGTCACCTTTCACTTGTGGTGGAAGAAAGCATCAGTGGCAACACCAGATCCAGTCTGTAG 1980
1921 gggtcaccttctacttgtgtggtggaagaaagcatcagtggaacaccagatccagctgtag 1980
QY 1981 AGCCATACACCAAGCAGCAGCTGTAACAACATGTGATTTGCTGAATATCATGTGGCTATA 2040
1981 agccatacaccaagcagcagctgtaacaacatgtgcatltgtcgaatacatatgagctata 2040
QY 2041 AGATCATGATGCGACCAACATCCGTTGCTTCACACTTGTCTTACCTGTACCCGACATTC 2100
2041 agatcatgatgcgaccacaacatccggttgccttcacacttgtcttacctatacccgaaatlc 2100
QY 2101 CCAAGGAGAGGCAATTTGGAAGTACTGTAGGCCCGAGACGCGACAGCACCCGAAACCCG 2160
2101 ccaaggagagagcaatttggaaagtactgtagcccgagagccaagagcaacccgaaagcg 2160
QY 2161 ACCCAAGTATGTCGCCCGGACCTGGAAGACCAAGTTCATGTGTGACACCAACGACCT 2220
2161 acccaagtatgtgcccgacccgtacatggaagaccaaagttcatctgtgtgacaccaaagcct 2220
QY 2221 GCAGCAATACCATTTGACCTGCGGATGTCGCCGACACTTTAGATTCAATTTGATGCAATTTG 2280
2221 gcagcaataccatttgacctgctgcgaltgtcccccgaacttgaattcatltgtagtltg 2280
QY 2281 GAAATTAACGGTGAAGCTGTGAGCCCTTACAGAGAGGCAATTTGAGTGCCTGACGTTTG 2340
2281 gaaatlaacggtgaagctgtgagcccttacagcagagagcgagcttggagctgtcaagctltg 2340
QY 2341 ACATGATCTGACCTCGGAGTGTGCTACCTCCCATGTGAGGAGCTGAAACGAGAGCT 2400

|||||
Db 2341 acatgatactgacctgagatgltgtaacctcccatgagagagctgaaccagaagt 2400
Oy 2401 GCAGAGAGTGAATTGAGACACCTGCCCCGTGCACCCCTAAGACGCCGAACCCATA 2460
Db 2401 gtagagagctgactgagagacacctgcccgtgctccacccctaagacagacccata 2460
Oy 2461 TCGTCTGAAACTCTTAACCTTTGTGCTCCAGATTTTTTTTTTAATTTCTACTTGTCT 2520
Db 2461 tgcgtcgaacctctaactcttggtlccagatttttttttlaatttccactctgtct 2520
Oy 2521 ACTTTGGGCATCTGCGACATTTTAAAGAGAGAAATGATGACGTGCGGTGATTAAC 2580
Db 2521 acccttggacaactcgggcaacttlttaaaagagaaatgagctggtggtgataaac 2580
Oy 2581 TGTATGTAAAGAGAGAGACCTGTGAGTCTGSGGATGGGGCTGAGAGCAGAGAGGC 2640
Db 2581 tgtatgtaagaagagagagacctctgagctctggtggtggtggtgagagagagagc 2640
Oy 2641 AAAGGGGAACACCT 2700
Db 2641 aaaggggaacacct 2700
Oy 2701 TGTGTAGACAGTGGCT 2760
Db 2701 tgtgtagacaaagtct 2760
Oy 2761 ACCCATTTCTGGGAACCT 2820
Db 2761 acccatctctgggaact 2820
Oy 2821 TGGGACTAAGCCAGAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2869
Db 2821 taggactaagccagagaggtctctctctctctctctctctctctctctctctctct 2869

RESULT 10
US-08-416-581B-7
; Sequence 7, Application US/08416581B
; Patent No. 5719042
; GENERAL INFORMATION:
; APPLICANT: Kishimoto, Tadamiatsu
; APPLICANT: Akira, Shizuo
; TITLE OF INVENTION: TRANSCRIPTION FACTOR APRT
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MIOM, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,581B
; FILING DATE: 04-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-65825/1994
; FILING DATE: 04-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Dean H.
; REGISTRATION NUMBER: 33,981
; REFERENCE/DOCKET NUMBER: 0-37891
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 2652 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-08-416-581B-7

Query Match 84.2%; Score 2415.6; DB 1; Length 2652;
Best local similarity 99.3%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

Oy 13 CAGCGCGCCAGTCCGGC--TCAGCCCGAGACAGTGCAGACCCCTGACTGCAGCAGAT 70
Db 201 CAGCGCGCCAGTCCGGCCTCAGCCCGAGACAGTGCAGACCCCTGACTGCAGCAGAT 260
Oy 71 GGCCTAGTGAACACACTGCAGCAGTGCAGACAGCTTCTGAAACAGTGCAGCAGCT 130
Db 261 GGCCTAGTGAACACACTGCAGCAGTGCAGACAGCTTCTGAAACAGTGCAGCAGCT 320
Oy 131 GTACATCGACAGCTTCCCATGAGCTGGGAGCTTCTGACACTTGTGATTGAGATGA 190
Db 321 GTACATCGACAGCTTCCCATGAGCTGGGAGCTTCTGACACTTGTGATTGAGATGA 380
Oy 191 AGACTGGCATATGCAGCCAGCAAGAGTCACATGCACGTTGGTGTTCATTAATCTTT 250
Db 381 AGACTGGCATATGCAGCCAGCAAGAGTCACATGCACGTTGGTGTTCATTAATCTTT 440
Oy 251 GGGTGAATTTGACACAGCAATATAGCCGATTCCTGCAAGAGTCCAAATGTCCTATGCA 310
Db 441 GGGTGAATTTGACACAGCAATATAGCCGATTCCTGCAAGAGTCCAAATGTCCTATGCA 500
Oy 311 CAACCTTCAAGAAATCAGACAGTCTTCTGCAAGACAGTATCTTGAGACAGCAATGGAAT 370
Db 501 CAACCTTCAAGAAATCAGACAGTCTTCTGCAAGACAGTATCTTGAGACAGCAATGGAAT 560
Oy 371 TGGCGGATCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 430
Db 561 TGGCGGATCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 620
Oy 431 GGCAGGCCAGCAAGGGGGGCGGAGGCCAACCACCCCAACAGCGCGGTGATGACAGAGA 490
Db 621 GGCAGGCCAGCAAGGGGGGCGGAGGCCAACCACCCCAACAGCGCGGTGATGACAGAGA 680
Oy 491 GCAGATCTTGGAGACACATCTTTCAGATGTCCGGAAGCAGATGTCAGATCTAGACAGA 550
Db 681 GCAGATCTTGGAGACACATCTTTCAGATGTCCGGAAGCAGATGTCAGATCTAGACAGA 740
Oy 551 AATGAAGTGTGGAGAACCTCCAGAGCAGCTTGTTCACACTACAAAACCTCAAGAG 610
Db 741 AATGAAGTGTGGAGAACCTCCAGAGCAGCTTGTTCACACTACAAAACCTCAAGAG 800
Oy 611 CCAGGAGACATGACAGATCTGAAATGAAACCAACAGTCTGTGACAGACAGAAATGCA 670
Db 801 CCAGGAGACATGACAGATCTGAAATGAAACCAACAGTCTGTGACAGACAGAAATGCA 860
Oy 671 GCAGTGGAAAGATGCTTCAAGCCCTGACACAGATGGGGAAGAAAGATGTGATAGCT 730
Db 861 GCAGTGGAAAGATGCTTCAAGCCCTGACACAGATGGGGAAGAAAGATGTGATAGCT 920
Oy 731 GGGGGGCTCTTGTGCAAGTGGAGTACGTGCAAGAACACTGACTGATGAAGAGTGGC 790
Db 921 GGGGGGCTCTTGTGCAAGTGGAGTACGTGCAAGAACACTGACTGATGAAGAGTGGC 980
Oy 791 TGACTGGAAGAGGGCGGCGACAGATGCGGTGCAATGGAGGCCCTCCCAACATGCTGCA 850
Db 981 TGACTGGAAGAGGGCGGCGACAGATGCGGTGCAATGGAGGCCCTCCCAACATGCTGCA 1040
Oy 851 CCGTGTGAAAATGATTAATCTTATAGCAATCTCAACTTCAAGACCCGCAACAAAT 910
Db 1041 CCGTGTGAAAATGATTAATCTTATAGCAATCTCAACTTCAAGACCCGCAACAAAT 1100
Oy 911 TAAGAACTGAGAGAGCTGCACAGCAAAAGTGTCTTCAAGAGGCGAGCCCTATGCTGCAGA 970


```

; ORIGINAL SOURCE:
; ORGANISM: Mouse
; TISSUE TYPE: Liver
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 259..2568
;
US-08-416-581B-8

```

Query Match	84.28;	Score 2415.6;	DB 1;	Length 2652;
Best Local Similarity	99.38;	Pred. No. 0;		
Matches 2436;	Conservative 0;	Mismatches 14;	Indels 2;	Gaps 1;

OY	13	CACCGCGCCAGTGGGCG--TGAGCCCGGAGACGTGAGAGACCCCTGACGTACACAGAT	70
Db	201	CAGCGCGCCAGTGGGCGCTCAGCGCCCGAGACAGTCGAGACCCCTGACTCAGCAGAGAT	260
OY	71	GGCTCAGTGAACCAACGCTGCAGCAGCTGAGACACAGCTACCTGAAGCAGCTGCACAGCT	130
Db	261	GGCTCAGTGAACCAACGCTGCAGCAGCTGAGACACAGCTACCTGAGCAGCTGCACAGCT	320
OY	131	GTACACCGACACGTTCCCTCATGAGAGCTGGCGGAGATTCTCTGGACCTTGGATTGAGAGTCA	190
Db	321	GTACACCGACACGTTCCCTCATGAGAGCTGGCGGAGATTCTCTGGACCTTGGATTGAGAGTCA	380
OY	191	AGACTGGGCATATGCAGCCAGCAGCAAGAAGTGCATATGCACATTGGTGTTCATTAATCTCTT	250
Db	381	AGACTGGGCATATGCAGCCAGCAGCAAGAAGTGCATATGCACATTGGTGTTCATTAATCTCTT	440
OY	251	GGGTGAATTTATCCAGCAATATATGCGGATTTCCTGCAGAGTGCATATGTCCTATACAGA	310
Db	441	GGGTGAATTTATCCAGCAATATATGCGGATTTCCTGCAGAGTGCATATGTCCTATACAGA	500
OY	311	CACCTTCGAAGAATTCAGCAGATTTCTGCAGACAGGATATCTTGAAGACCAATGGAAAT	370
Db	501	CACCTTCGAAGAATTCAGCAGATTTCTGCAGACAGGATATCTTGAAGACCAATGGAAAT	560
OY	371	TGCCGGATTCGTGGCCCGGATGCTGTGGGAAGAGTCTGCCTCTCCACAGCGGACGCAC	430
Db	561	TGCCGGATTCGTGGCCCGGATGCTGTGGGAAGAGTCTGCCTCTCCACAGCGGACGCAC	620
OY	431	GGCAGCCAGCAAGGGGGCCAGGCGCAACCCACACAGCCGCGGTATGTACAGAGAAACA	490
Db	621	GGCAGCCAGCAAGGGGGCCAGGCGCAACCCACACAGCTGCCGTATGTACAGAGAAACA	680
OY	491	GCAGATGTTTGGAGCAGCATCTTCAGAGATGTCCGGAAAGCAGTGCAGGATCTAGAACAAGA	550
Db	681	GCAGATGTTTGGAGCAGCATCTTCAGAGATGTCCGGAAAGCAGTGCAGGATCTAGAACAAGA	740
OY	551	AATGAAGTGGTGGAGAAACCTCCAGACAGCATTTGATTTCACTACCAAAACCCCTCAAGAG	610
Db	741	AATGAAGTGGTGGAGAAACCTCCAGACAGCATTTGATTTCACTACCAAAACCCCTCAAGAG	800
OY	611	CCAAAGAGACATATGCAGAGATCTGAATGGAAACAACCGTGTGTACACAGACAGAAAGATCA	670
Db	801	CCAAAGAGACATATGCAGAGATCTGAATGGAAACAACCGTGTGTACACAGACAGAAAGATCA	860
OY	671	GCAGTGGAAAGATGTCTCACAGCCCTGGACCAAGATGGCGGAGAACCATTTGAGTGAAGACT	730
Db	861	GCAGTGGAAAGATGTCTCACAGCCCTGGACCAAGATGGCGGAGAACCATTTGAGTGAAGACT	920
OY	731	GGCGGGGCTCTTGTCAAGCAATGGAGTACGTGCAGGAAGACACTGACTGATGGAAGAGCTGGC	790
Db	921	GGCGGGGCTCTTGTCAAGCAATGGAGTACGTGCAGGAAGACACTGACTGATGGAAGAGCTGGC	980
OY	791	TGACGTGGAAGAGGGGGCCAGAGATCGGTGTGCATCGGAGGACCTCCCAACATCTGGCTGGA	850
Db	981	TGACGTGGAAGAGGGGGCCAGAGATCGGTGTGCATCGGAGGACCTCCCAACATCTGGCTGGA	1040
OY	851	CCGTCTGGAAAACTGATTAACCTTCATTAGCAGAACTCACTTCAGACCCCGCAACAAT	910
Db	1041	CCGTCTGGAAAACTGATTAACCTTCATTAGCAGAACTCACTTCAGACCCCGCAACAAT	1100

QY	911	TAAGAAATCGAGAGCTGCAGACGAAAGTGTCTACAAAGGCGCACCTATCTGTCAGCA	970
DB	1101	TAAGAAACTGGAGAGAGCTGCAGCAGAAAGTGTCTTCAAGAGGCGCACCTATCTGTCAGCA	1160
QY	971	CCGGCCATCTCTGGAGAGAGCATCTGGAGAGCTTGCACAACTTAATGAAGAGTGCCT	1030
DB	1161	CCGGCCATCTCTGGAGAGAGGATCTGGAGAGCTTGCACAACTTAATGAAGAGTGCCT	1220
QY	1031	CGTGTGGAGCGGAGCCCTGCATATGCCATATGACCCCGGACCGGCGCTTAGTATCATAGAC	1090
DB	1221	CGTGTGAGCGGAGCCCTGCATATGCCATATGACCCCGGACCGGCGCTTAGTATCATAGAC	1280
QY	1091	TGGTGTCCAGTTTACACAGAAATGTCAGGTTGCTGGTCAATTTCTCGAGTTGAATTTATCA	1150
DB	1281	TGGTGTCCAGTTTACACAGAAATGTCAGGTTGCTGGTCAATTTCTCGAGTTGAATTTATCA	1340
QY	1151	GCTTAATTTAAAGTGTGCATATGAAGACTTGGGGATGTTGTGCCCTCAGAGGTC	1210
DB	1341	GCTTAATTTAAAGTGTGCATATGAAGACTTGGGGATGTTGTGCCCTCAGAGGTC	1400
QY	1211	TGCGAATTTTAACTTCTGGGCGACGAACACAAAAGTATGAACATATGAGAGACTTAACAA	1270
DB	1401	TGCGAATTTTAACTTCTGGGCGACGAACACAAAAGTATGAACATATGAGAGACTTAACAA	1460
QY	1271	CGGCGAGCTGTCTGCAGAGATTCAAGACACCTGCAGCTTACGGAGAGCAGAGATGGGAATG	1330
DB	1461	CGGCGAGCTGTCTGCAGAGATTCAAGACACCTGCAGCTTACGGAGAGCAGAGATGGGAATG	1520
QY	1331	AGGCGCGCCAAATTTGTATGCTCTTGTATCTGTACTGAGAGAGCTGCACCTGATACCTT	1390
DB	1521	AGGCGCGCCAAATTTGTATGCTCTTGTATCTGTACTGAGAGAGCTGCACCTGATACCTT	1580
QY	1391	CGAGCATAGGTTGACCAACAAGGCGCTCAAGATTTGACCTAGAGAGCCACCTCGTCCGCACT	1450
DB	1581	CGAGCATAGGTTGACCAACAAGGCGCTCAAGATTTGACCTAGAGAGCCACCTCGTCCGCACT	1640
QY	1451	TGTGTGTATCTCCAAACATCTGTGCAGATGCCAAATGCTTGGGCAATCTGTGTATAA	1510
DB	1641	TGTGTGTATCTCCAAACATCTGTGCAGATGCCAAATGCTTGGGCAATCTGTGTATAA	1700
QY	1511	CATGCTGCACCAATTAACCCCAAGAACGTGACCTCTTCACTAAAGCGCCCAATTGGAACTG	1570
DB	1701	CATGCTGCACCAATTAACCCCAAGAACGTGACCTCTTCACTAAAGCGCCCAATTGGAACTG	1760
QY	1571	GGACCAAGTGGCCGAGGTGCTCAGCTTGGCATTTCTGTCACACACCAAGCGAGGCTGAG	1630
DB	1761	GGACCAAGTGGCCGAGGTGCTCAGCTTGGCATTTCTGTCACACACCAAGCGAGGCTGAG	1820
QY	1631	CATGCAACCAATTTGACAAAGCTGGGTGAGAAGCTCTTAGGGCTTGGTGTGAACATCTACAG	1690
DB	1821	CATGCAACCAATTTGACAAAGCTGGGTGAGAAGCTCTTAGGGCTTGGTGTGAACATCTACAG	1880
QY	1691	GTTGTAGATATCATATGGCTTAATTTCTGCAAGAAGAACATGCGTGGGCAAGGCTTCTCTT	1750
DB	1881	GTTGTAGATATCATATGGCTTAATTTCTGCAAGAAGAACATGCGTGGGCAAGGCTTCTCTT	1940
QY	1751	CTGGCACTCTGCTAGCAATATCATGCACTTTGGAAGAAAGTATATCTTGGCCCTTTGGAA	1810
DB	1941	CTGGCACTCTGCTAGCAATATCATGCACTTTGGAAGAAAGTATATCTTGGCCCTTTGGAA	2000
QY	1811	TGAAGGATCATCATGGGTTTATCATCACAGAGCGGAGCGGGCCATCTTAAGACAAA	1870
DB	2001	TGAAGGATCATCATGGGTTTATCATCACAGAGCGGAGCGGGCCATCTTAAGACAAA	2060
QY	1871	GGCCCGGGGACCTTCTACTGCGCTTCAGCGAGAGCAGCAAAAGAGGGGCTCACTT	1930
DB	2061	GGCCCGGGGACCTTCTACTGCGCTTCAGCGAGAGCAGCAAAAGAGGGGCTCACTT	2120
QY	1931	CACCTTGGTGAAGAAAGACATCACTGGCAGACCCAGATATCGATGTAGAGCCATACAC	1990
DB	2121	CACCTTGGTGTGAAGAAAGACATCACTGGCAGACCCAGATATCGATGTAGAGCCATACAC	2180
QY	1991	CAGAGAGCAGCTGAACAACATGTCATTTGCTGAAATCATCATAGGGCTATATAGATATGGA	2050


```

Db 961 TTTCTGTGGGCGGCGAGCCCTGCATGCCATGACCCGGGCCCTTAGTATCAAG 1020
Oy 1089 ACTGCTGCCAGTTTACACGAAGTCAGTGTGTCATTAATTCTGAGTTGAATTA 1148
Db 1021 ACTGCTGCCAGTTTACACGAAGTCAGTGTGTCATTAATTCTGAGTTGAATTA 1080
Oy 1149 CAGCTTAAATTAAGTGTGATTAAGACTCTGGGGATGTTGCTGCCCTCAGAGGG 1208
Db 1081 CAGCTTAAATTAAGTGTGATTAAGACTCTGGGGATGTTGCTGCCCTCAGAGGG 1140
Oy 1209 TCTCGAAATTAATTAATTCCTGGGCGACGACAAAGATGATGAGAGAGTGAAC 1268
Db 1141 TCTCGAAATTAATTAATTCCTGGGCGACGACAAAGATGATGAGAGAGTGAAC 1200
Oy 1269 AACGGACCCCTGTCTGAGAGTTCAGACCCCTGACCCCTTAGAGGACAGATGTGGAA 1328
Db 1201 AACGGACCCCTGTCTGAGAGTTCAGACCCCTGACCCCTTAGAGGACAGATGTGGAA 1260
Oy 1329 GGAGGCCCTGCCAATGTGATGCTCTTGTATCGTGAATGAGAGTGCACCTGATCAC 1388
Db 1261 GGAGGCCCTGCCAATGTGATGCTCTTGTATCGTGAATGAGAGTGCACCTGATCAC 1320
Oy 1389 TTCGAGACTGAGTGTACACCAAGGCTCAAGATTGACTAGAGACCCACTCTTGCCA 1448
Db 1321 TTCGAGACTGAGTGTACACCAAGGCTCAAGATTGACTAGAGACCCACTCTTGCCA 1380
Oy 1449 GTTGTGTGATCTCCACATCTGTACATGCCAAATGCTTGGGATCAATCTGTGTAT 1508
Db 1381 GTTGTGTGATCTCCACATCTGTACATGCCAAATGCTTGGGATCAATCTGTGTAT 1440
Oy 1509 AACATGCTGACCAATTAACCCCAAGACGTGAACCTTCTCACTAAGCCGCAATTGGAAC 1568
Db 1441 AACATGCTGACCAATTAACCCCAAGACGTGAACCTTCTCACTAAGCCGCAATTGGAAC 1500
Oy 1569 TGGACCAAGTGGCGGAGGCTCTAGCTGGCATTTCTGTCACACCAAGCGAGGCTG 1628
Db 1501 TGGACCAAGTGGCGGAGGCTCTAGCTGGCATTTCTGTCACACCAAGCGAGGCTG 1560
Oy 1629 AGCATCGAGCTGACCAAGCCTGGCTGAGAAGCTCTAGGAGGCTGGGTGAAGTACTCA 1688
Db 1561 AGCATCGAGCTGACCAAGCCTGGCTGAGAAGCTCTAGGAGGCTGGGTGAAGTACTCA 1620
Oy 1689 GGGTGTGATTCACATGGGCTAAATTCGCAAGAAAACATGGCTGGCAAGGGCTTCTCC 1748
Db 1621 GGGTGTGATTCACATGGGCTAAATTCGCAAGAAAACATGGCTGGCAAGGGCTTCTCC 1680
Oy 1749 TTTGGGCTGTGGCTAGACAATATCATGACCTGTGAAAAGTATATCTTGGCCCTTGG 1808
Db 1681 TTTGGGCTGTGGCTAGACAATATCATGACCTGTGAAAAGTATATCTTGGCCCTTGG 1740
Oy 1809 AATGAAGGTCATCATGAGTTTCATCAGCAAGAGCGGAGCGGACCTCTAAGCACA 1868
Db 1741 AATGAAGGTCATCATGAGTTTCATCAGCAAGAGCGGAGCGGACCTCTAAGCACA 1800
Oy 1869 AAGCCCCCGGCGACCTTCTTACTGGCTTCAGCGAGACAGCAAGAAGAGGGTCACT 1928
Db 1801 AAGCCCCCGGCGACCTTCTTACTGGCTTCAGCGAGACAGCAAGAAGAGGGTCACT 1860
Oy 1929 TTTCACTTGGTGAAGAAAGATCAGTGGCAAGACCCAGATTCAGTGTGAGAGCAATAC 1988
Db 1861 TTTCACTTGGTGAAGAAAGATCAGTGGCAAGACCCAGATTCAGTGTGAGAGCAATAC 1920
Oy 1989 ACCAAGCAGCAGTCGAACACATGTCATTCCTGAATCATCATGCGCTATTAAGTCATG 2048
Db 1921 ACCAAGCAGCAGTCGAACACATGTCATTCCTGAATCATCATGCGCTATTAAGTCATG 1980
Oy 2049 GATGCGACCAACATCTGCTGCTCAGTGTGCTTCACTCTTACCCCGAATTTCCCAAGAG 2108
Db 1981 GATGCGACCAACATCTGCTGCTCAGTGTGCTTCACTCTTACCCCGAATTTCCCAAGAG 2040
Oy 2109 GAGGCAATTTGGAAGTACTGTAGCGCGAGAGCCAGAGACCCCGAAGCCGACCAAGT 2168

```

```

Db 2041 GAGGCAATTTGCAAGTACTGTAGGCGCCGAGAGCCAGACACCCCGAAGCCGACCAAGT 2100
Oy 2169 AGTCTGCCCGCTACCTGAGAACCAAGTTCAATCTGTGTGACAGCAGACCTGAGCAAT 2228
Db 2101 AGTCTGCCCGCTACCTGAGAACCAAGTTCAATCTGTGTGACAGCAGACCTGAGCAAT 2160
Oy 2229 ACCATTTGACCTGCCGATGTCTCCCGGACACTTACATTTGATGATGCAATTTGAAATTAAC 2288
Db 2161 ACCATTTGACCTGCCGATGTCTCCCGGACACTTACATTTGATGATGCAATTTGAAATTAAC 2220
Oy 2289 GGTGAAGTGTGAGCCCTCAGCAGAGAGGCGATTTGAGTGCCTCAGCTTTGAGATGAT 2348
Db 2221 GGTGAAGTGTGAGCCCTCAGCAGAGAGGCGATTTGAGTGCCTCAGCTTTGAGATGAT 2280
Oy 2349 CTGACCTGAGAGTGTCTACTCTCCCATG 2378
Db 2281 CTGACCTGAGAGTGTCTACTCTCCCATG 2310

RESULT 13
US-08-416-581B-3
: Sequence 3, Application US/08416581B
: Patent No. 5719042
: GENERAL INFORMATION:
: APPLICANT: Kishimoto, Tadamiatsu
: APPLICANT: Akira, Shizuo
: TITLE OF INVENTION: TRANSCRIPTION FACTOR APPE
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SUGHRUE, MIION, ZINN, MACBEAR & SEAS
: STREET: 2100 Pennsylvania Avenue, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/416,581B
: FILING DATE: 04-Apr-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-65825/1994
: FILING DATE: 04-Apr-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Nakamura, Dean H.
: REGISTRATION NUMBER: 33,981
: REFERENCE/DOCKET NUMBER: O-37891
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)293-7060
: TELEFAX: (202)293-7860
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2787 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
US-08-416-581B-3

Query Match 74.7%; Score 2144.2; DB 1; Length 2787;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 2305; Conservative 0; Mismatches 223; Indels 7; Gaps 2;
Oy 61 GCACGAGATGGCTCAGTGAACCACTGACAGCAGCTGACAGCAGCTTACCTGAAGCAGC 120
Db 213 GAAACAGGATGGCCCAATGGAATCAGCTACAGCAGCTTGAACACAGGATCCTGAGAGCAGC 272
Oy 121 TGCAACAGCTGTACAGGACACGTTTCCCATGAGAGTGTGCGGCAATTCCTGTGACACTTGA 180

```


Db	1233	TCATCAAGACCGGGCTCCAGTTCTACTACTAAAGTCCAGGTTGCTGGTCAAGTTCCCTGA	1292
Qy	1141	TGAATTTATCAGCTTTAAATTAATTAAGTGTGCAATTGAATTAACACTCTGGGGATGTTGCTGCC	1200
Db	1293	TGAATTTATCAGCTTTAAATTAATTAAGTGTGCAATTGAATTAACACTCTGGGGATGTTGCTGCC	1352
Qy	1201	TCAGAGGCTCTCGGAAATTTAAATTTCAATTTCTGGGACGCAACCAAAACTGATGAACTGGAGG	1260
Db	1353	TCACAGAGATCCCGGAAATTTAAATTTCAATTTCTGGGACGCAACCAAAACTGATGAACTGGAGG	1412
Qy	1261	AGTGTAAACAACGCGAGCGCTGTCTGCAGAGTTGCAAGCACCTGACCCCTTAGAGGAGCAGAGAT	1320
Db	1413	AATCCAAACAACGCGAGCGCTGTCTGCAGAGTTGCAACCACTTAGACCTGAGGAGCAGAGAT	1472
Qy	1321	GTGGGAATGCAAGGCGCTGGCCAATTTGGATGGCTCTTGAATCTGTGACTGAGAGACTGCAC	1380
Db	1473	GTGGGAATGCGGGGCGCGAGCCAAATTTGGATGGCTCTTGAATTTGACTGAGAGAGCTGCAC	1532
Qy	1381	TGATACACCTTTCGAGACTGAGAGTGTACCAACAAGGCGCTCAAGATTTACCTAGAGACCACT	1440
Db	1533	TGATACACCTTTGAGACCGGAGGTGTATCCACAAGGTCTCAAGATTTACCTAGAGACCACT	1592
Qy	1441	CCTTGCCAGTTGTGTGATATCTCCACATCTGTCTCAGATGTCCAAACTCTTGGGATCAATCC	1500
Db	1593	CCTTGTAGTGTGTGTATCTCCACATCTGTCTCAGATGTCCAAACTGCTGGGGGTCTCAATCC	1652
Qy	1501	TGTGTATTAACATGCTGTACCAATTAACCCCAAGAAGCTGAATCTTCACTAAGCGGCCAA	1560
Db	1653	TGTGTATACACATGCTTACCAACAATTCACAGAATGTGAATCTTCTTCACTAAGCGGCCAA	1712
Qy	1561	TTGCAACCTGGGAGCAAGTGGCGGAGGTGCTCAGCTGGAGTTCTGTCCACACCAACAGC	1620
Db	1713	TTGGAACCTGGGAGCAAGTGGCGGAGGTGCTCAGCTGGAGTTCTGTCCACACCAACAGC	1772
Qy	1621	GAGGCTGTAGATGTGACACACTGTACAACAGTGGGCTGAGAAAGCTCTAGGGCGCTGGTGGA	1680
Db	1773	GAGGCTGTAGATGTGACACACTGTACAACAGTGGGCTGAGAAAGCTCTAGGGCGCTGGTGGA	1832
Qy	1681	ACTACTCAGGCTGTCACATACATGTGGCTTAAATTTGCAAAAGAAACAATGTGCTGCAAGG	1740
Db	1833	ACTACTCAGGCTGTCACATACATGTGGCTTAAATTTGCAAAAGAAACAATGTGCTGCAAGG	1892
Qy	1741	GCTTCTCCTTGGGCTGTGGCTAGACAAATATCATGCACTTGTGAAAAAGTATATCTTGG	1800
Db	1893	GCTTCTCCTTGGGCTGTGGCTAGACAAATATCATGCACTTGTGAAAAAGTATATCTTGG	1952
Qy	1801	CCCTTTGGAATGAAGGGTATCATATGAGTTTCTCAGCAAGAAGCGGAGCGGCATCC	1860
Db	1953	CCCTTTGGAATGAAGGGTATCATATGAGTTTCTCAGCAAGAAGCGGAGCGGCATCT	2012
Qy	1861	TAAACCAAAAGCCCCGGGACCTTCTACTGCGCTTCAAGCGAGACCAAAAGAGAG	1920
Db	2013	TGAGACTTAACCCCCCGGAGACCTTCTGCTGCGCTTCAAGCGAGCAAAAGAGAGAG	2072
Qy	1921	GGGTACTTTTACTTGGGTGGTGAAGAGACATCGTGGCAAGACCCAGATCCAGTCTGAG	1980
Db	2073	GGGTACTTTTACTTGGGTGGTGGAGAGAGACATCGCGGTAAAGCCAGATCCAGTCTGAG	2132
Qy	1981	AGCCATACACCAGACAGACTGAAACAATGTCAATTTGCTGAATTCATCATGGGCTATA	2040
Db	2133	AACCATACACAAGCAGCAGGTGAAACAATGTCAATTTGCTGAATTCATCATGGGCTATA	2192
Qy	2041	AGATCATGTGATGCAACCAACATCTGTGTCTCCACTTGTGTACCTGTACCCGACATTC	2100
Db	2193	AGATCATGTGATGCAACCAACATCTGTGTGTCTCCACTTGTGTACCTGTACCTGACATTC	2252
Qy	2101	CCAAGGAGAGGAGATTGGGAAAGTACTGTAGGCCAGAGAGCCAGCCGAGAGCGG	2160
Db	2253	CCAAGGAGAGGAGATTGGGAAAGTATTGTGGGCAAGAGCCAGAGCATCTCTGAAGCTG	2312
Qy	2161	ACCCAGTAGTGTGCCCCGTACTGTGAAGCAAGTTCAATCTGTGTGACACCAACGACCT	2220

Db	2313	ACCCAGGTAGCGCTGCCCATACCTGAAGACCAAGTTTATCTGTGTGACACCAACGACCT	2372
QY	2221	GCACGAATACCATTTGACCTCCGATGTCCCGCCGACCTTAGATTCAATGATGCAATTGG	2280
Db	2373	GCACGAATACCATTTGACCTCCGATGTCCCGCCGCTTTAATTTGATTTGATGCAAGTTTG	2432
QY	2281	GAATTAACGCGTGAAGTGTCTGAGCCCTCAGCAGAGGAGCATTTGATGCGTCACGTTTG	2340
Db	2433	GAATTAATGTGTGAAGTGTCTGCAACCCTCAGCAGAGGAGCATTTGATGCGTCACGTTTG	2492
QY	2341	ACATGATCTGACCTCGGAGTGTCTACCTGCCCCATGTTGAGGAGCTGAAACAGACACT	2400
Db	2493	ACATGAGATTGACTCGGAGTGTGCGTACACCTCCCATGTTGAGGAGCTGAAACAGACACT	2552
QY	2401	GC---ACAGACGTGACTTAGACACCTGCGCCGTGCTCCACCCTTAAGCAGCCGCAACC	2456
Db	2553	GCAGAAAGATATCGACTGAGCGCGCTACCTTCATTTGCCACCCCTCACACACCCAAACCC	2612
QY	2457	CATATCGCTGAAMCCTCTCACTTTGGTGTCCAGATTTTTTTTTTATTTTCCACTTC	2516
Db	2613	CAGATCATCTTAAACCTCTCACTTTGGTGTCCAGA---TTTTTTTTTAAATCTCCACTTC	2669
QY	2517	TGCTATCTTTGGCAATCTGGGACCTTTTAAAGAGAGAAATGAGTGTGGTGAT	2576
Db	2670	TGCTATCTTTTAGCAATCTGGGACCTTTTAAATGAGAAATGAGTGTGGTGAT	2729
QY	2577	AAACGTATGTATAA	2591
Db	2730	CTGCTTTATCTTAAA	2744

RESULT 15

US-09-288-461-1
; Sequence 1, Application US/09288461

; Patent NO. 6159694
; GENERAL INFORMATION:

```

; APPLICANT: Karras, James G.
; TITLE OF INVENTION: Antisense Oligonucleotide Modulation of STAT3

```

```

; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0338

```

; CURRENT APPLICATION NUMBER: US/
; CURRENT FILING DATE: 1999-04-08

```
; NUMBER OF SEQ ID NOS: 107
;
; SOFTWARE: PatentIn Ver. 2.0
```

```

; SEQ ID NO 1
; LENGTH: 2787

```

```

; TYPE: DNA
; ORGANISM: Homo Sapiens

```

```

; FEATURE:
; NAME/KEY: CDS
;

```

; LOCATION: (221)..(2533)
; PUBLICATION INFORMATION:
;; JOURNAL: Cell
; VOLUME: 77
; 1994

ISSUE: 1
PAGES: 63-71

; DATE: 1994-0
; DATABASE ACCO
; DATA

DATE: 1994-12-31
US-09-288-461-1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Query Match	Score	DB	Length
Best Local Similarity	90.98%	Pred. No. 0;	

Matches	2305;	Conservative	0;	Mismatches	223;	Inserts	1;	Gaps	130
---------	-------	--------------	----	------------	------	---------	----	------	-----

61 GCACGACGATGGCTCAGTGGACACGCTGACGACAGCTGACACACACCTATCCTGATGACAGC 120

[illegible][illegible]

OY	181	TTGAGACTCAAGACTGGGCATATGACAGCCACGAAGAAGTCACATGCCAGTTGGTGTTC	240
Db	333	ctgagaaatcaagaattcgtggcatalctgctgcgcagcaaaagaaatcacatctgcactcttggtcttc	392
OY	241	ATATATCTCTTGGGGTGAATTTGACCAGCATATATGAGCCGATCTCTGGAAAGTGCATATGCC	300
Db	393	ataaactcctcgtggagagaaattgcaccgaagatatagcgcctcttcgcgaagatctgaattcttc	452
OY	301	TCTATCAGCACAACTTTCGGAAGATCAAGCAGTCTTCTGCAGAGCAGGATCTTGTGAAGAC	360
Db	453	ctcatcagcaacaatctcagaagaatcaagagttctctcagagcagtaatctctgagaagc	512
OY	361	CAATGGAATTTGGCCCGGATGTGTGGCCCGATGCTCGTGGGAAGAAGTCTGGCCCTCCACGA	420
Db	513	caatgagagatttccccggaattctgtgcccggctgcctctgtggagaatcaacgctctctacga	572
OY	421	CGCAGCGCCAGCGCAGCCACGAAGGGGGCCAGGCCAACCCACCAACAGCGCCGCTGATGA	480
Db	573	ctgcagcccaatctgcggcccgcaagggggccaggccaaacccccaagaagagcgtggtgga	632
OY	481	CAGAGAAACACAGATGTTGTGAGCAGCATCTTACAGATGTCCGGAAGCAGATGCAGATC	540
Db	633	cggagaaagcagcagaaatgcctgtagcagacacttcagaaatgtccggaagagatgctcagagtc	692
OY	541	TAGAACAGAAAATTAAGGTGGTGAGAACCTCCAGAGCACATTTATTTCACTCAATAA	600
Db	693	tagaacaagaaaatgaaagtggttagagaatctccagaaatgactttagtcttaactataaa	752
OY	601	CCCTCAAGAGCCCAAGAGACATGCAGGATCTGAAATGGAACAACCAAGTCTGTGACCAGAC	660
Db	753	cccccaagagatcaagygagacatgcgaagatcttgaaatgaaacaacagatcagtgagaccagc	812
OY	661	AGAAGATGACGACCTCTGSAACACATGCTCACGCCCTCGACAGATGGGGAAACAACATTG	720
Db	813	agaagaatgcagcaacatgcggaacagaatgcctcagctgcctggaaccagaatgcggaagaatcag	872
OY	721	TGAGTGACCTGGCGGGCTTGTGTACAGCAATGAGTACTGTGCAGAAAGACACTGATGTAG	780
Db	873	tgagtgcgctgcggggcctcttgctcagcgaatgagtagctgcgaagaaactctcacgagc	932
OY	781	AAGAGCTGGCTGACTGGAAGAAGCGCGCACAGATGCGCTGTGACTCGSAGGCCCTCCCAACA	840
Db	933	agagagctgcgctgcgactctggaagaagcgcggaacaagaatctgcctgcgaatctggaagcccgcaaca	992
OY	841	TCTGCTCGAGCCGCTGTGGAAGAACCTGCATTACTTACTTACCACAATCTCAACTTCAGACC	900
Db	993	ctctgcctcagatcgcgctagaanaaacatgataacgtcatctagcagaatctcaactrcagagacc	1052
OY	901	GCCAACAATTTAGAAACTGAGAGAGCTGCAGCAGAAAGTGTCTTACANAGGCGACCCCTA	960
Db	1053	gtcacacaataatgaagaacatctgagagagatctgcacaaaagaatctctcacaagaagggcccaca	1112
OY	961	TCTGTCGACACCGGCCCATCTCTGAGAGAGAGATCTGTGAGCTGTTCGAAACTTAATGA	1020
Db	1113	ctgtacagcaacccgcgcgaatgcctggaagagagatactgtagagctgtctcagaaacttaataga	1172
OY	1021	AGAGTGCCTTCTGTGTGTGAGCGCGCAGCCCTGCATGCGCCCATGACCCCGAGCCGCCCTT	1080
Db	1173	aaagtctcctcttggtgtgtagcgcgacgcccctgcgatcgcacatcatctcgtacgcgccccctcg	1232
OY	1081	TCATCAGACAGCGGTGTCACATTTACACGAAGAAGTCAAGTGTGTGTCAAATTTCTGTAGT	1140
Db	1233	tcatacaagaccgggtcccaatctcaactcaataaagtcgaagttctgtgtcgaagtctccctgagt	1292
OY	1141	TGAATTTACAGCTTTAAAATTTAAAGTGTGCATTGTATTAAGACTCTGGGATGTGCTGCC	1200
Db	1293	tgaattcatcagcttaaatatlaaagtgtgatcttagtacaagaatccttggtgagcgtc	1352
OY	1201	TCCAGAGGCTCTCGGAATTTTAAATTTCTGCGCAGCAACAACAAAAGTGAATGAATGAGAGC	1260
Db	1353	tcgaagagatcccggaatcttaaatctcttggtgcacaaacaaaagctgtagtaactctgaaag	1412

QY	1261	AGTCTAAATACGGCAGCCTGTCTTCACAGTTTCAGACACCTGACACCTTATAGGAGCAGACAT	13260
Db	1413	aatcaacaacaagcgagccctctctgcagaaattcaacaacattgcacctgcgaagggagcagaat	14727
QY	1331	GTGGAAATGAGGCCCGGACCAATGTGTATCCCTCTTTCATGCTGACTGTAGAGCTGCACCC	1380
Db	1473	gtgggaatagggggcccagagccaatttgatgctctccctgattgtgactgaggagctgcac	15333
QY	1381	TGATACCTTTCGAGACTGAGCTGTACCAACCAAGCCTCAAGATTGTGACTAGACCCACT	14400
Db	1533	tgatcacctcttgagaccgaggtgatccaagaagctcagaatttgacatggaccatagagaccac	15929
QY	1441	CGTTGCCAGTTGTGTGATATCCCAACATCTGTGATATGCGCAATGTGTGGATATCAATC	15000
Db	1593	ccgtgcacgtctgtgtgattcctccaacatcgtctagaatgcacaatgctcgtggtccatccc	16550
QY	1501	TGTGTATTAACATCTGTCAGCAATTAACCCCAAGACCTGAACTTCTTCACTAAGCCGCCAA	15600
Db	1653	tgtgtataacaatgctgaccacaacatcccaagaatgtagaactctctcactaaggccgcaa	17120
QY	1561	TTGGAACCTGGGACCAAGTGGCCGAGTGTCTCAGCTTGGCAGTTCTGTCTCCACCACCAAGC	16220
Db	1713	ttggaaccttggaaccagtggtgcgaaggtgtctcagcttggcagttctgtctccaccaccagc	17727
QY	1621	GAGGCTCAGCATGAGACAGCTGCACAACGCTGGCTAGAAAGCTCTAAGGGCTGTGTGA	16800
Db	1773	gaggctcagcatgagacagctgcacaacgctggctagaagctctaaaggctgtgtga	18330
QY	1681	ACTACTAGGCTGTTCACATCATATGGGCTAAATTCTGCAGAAAGAAACATGCTGCACAGC	17400
Db	1833	actactaggctgttccatcatatcattggtcttaactcttcgacaagaacaatgctgcgaag	18929
QY	1741	GCTTCCTGTCTCGTGGCTGTAGACATTAATTCACACCTTGTGAAGATTAATCTTGG	18000
Db	1893	gcttccctgtctggctgttagacattaatctcacaccttgtgaagattaatcttgg	19550
QY	1801	CCCTTTGCAATGAAGGCTAGATCATAGGTTTCATCAGACAGACGCGGAGCGGCATCC	18600
Db	1953	ccctttgcaatgaaggctagatcataggtttcatcagcaagagcggtgagcgcatct	20120
QY	1861	TAAACCAAAACCCCCCGGCACCTCTCTACTGGGCTTCACCGAGACACCAAGAAAGAG	19220
Db	2013	tgaagcaaaaccccccgacacctctcgtctggtctcagttgaagcaagcaaaagagag	20727
QY	1921	GGGTCACTTACTTGGGTGGAAAGACATCAGTGGCAAGACCCAGTCCAGTGTGAG	19800
Db	2073	gggtcacttacttgggtggaaagacatcagtggtgcaagacccagtcagtgtag	21320
QY	1961	AGCCATTAATACAGCAGCTGACACACATGTTCATTTGCTGAATCATATGCGCTATA	20400
Db	2133	aaccttaataacagcagcagctgaacaacatgctcatctgtatatacatcatggtcata	21920
QY	2041	AGATTAAGCAATGCGACCAACATCTCGTGTCTCAGCTTGTCTACCTTCACCCGCATTC	21000
Db	2193	agatttaagcaatgcgacccaacatctcgtgtgtctcagcttgtctacacctgatactcctgacatc	22520
QY	2101	CCAAAGACAGGCAATTTTGAAAGTACTGTAGGCCGAGACCCAGACACCCCGAAGCG	21600
Db	2253	ccaaagacaggcaattttgaaagtactgttaggccgagacccagacaccccgaaagcg	23120
QY	2161	ACCCAGGTAGTGCCTGCCGACCTGAAGACCAAGTTCACTGTGTGACACCAAGCACT	22200
Db	2313	accaggtagtgctgcctgccgacctgaagacccaagttcatctgtggaaccaagacct	23720
QY	2221	GCAGCAATACCAATTGACCTCCGATGTCCCCCGCACTTTAGATTTCATGTATGACGTTTG	22800
Db	2373	gcagcaataccattgacctctgcgattgcccccgcgtctttagattcatgatgacgttg	24320
QY	2281	GAATTAACGGTGAAGGTGCTGACCTCTACAGACAGAGAGGCAAGTTAGTGTGCTCAGCTTG	23400
Db	2433	gaatataagtgaaaggtgctgaaaccttcagcaagaggttagttagtccctcacctctt	24920
QY	2341	ACATGATCTGCACCTCGAGTGTGCTACTCCCCCATGTGAGAGACTGAACACAACT	24000


```
Db 2493 ||||||| ||||||||| ||||||||| ||||||||| ||||||| 2552
      acatggagttgacctcggagtgcgctacccctcccaatgagagctgagaaacggaagct
QY 2401 GC---AGACAGCGTGACTTGAGACACCTGGCCCGTGCTCCACCCCTAAGCAGCCGAACCC 2456
      || || || || | | | | | | | | | | | | | | | | | | | | | |
Db 2553 gcagaaagatacgaactgagcgccctacctcgtcatctcgccacccctcacacagccaaccc 2612
      || || || || || || || || || || || || || || || || || || || ||
QY 2457 CATATCGTCTGAAACTTCCTAAGTTGTGTTCCAGATTTTTTTTAAATTCCTACTTC 2516
      || || || || || || || || || || || || || || || || || || || ||
Db 2613 cagatcacctgaaactactacttctggtccaga---tlttlttaalcctcctctc 2669
      || || || || || || || || || || || || || || || || || || || ||
QY 2517 TCGTATCTTTGGGCAATCTGGGCACTTTTAAAAAGACAGAAATGAGTGAATGTGGTGAAT 2576
      || || || || || || || || || || || || || || || || || || || ||
Db 2670 tgcatacttcgagcaatcggcaactttaaataagagaatcgaatgaaatggtgagc 2729
      || || || || || || || || || || || || || || || || || || || ||
QY 2577 AAACGTATTATGTAAA 2591
      || || || || ||
Db 2730 ctgcttltatcctaaa 2744
```

Search completed: March 19, 2002, 15:19:34
Job time: 4716 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2002, 14:00:57 ; Search time 1732.41 Seconds
(without alignments)
17795.810 Million cell updates/sec

Title: US-08-212-185-11

Perfect score: 2869
Sequence: 1 GCCGCCACCGCCAGCCGCG.....AATTAAAAAAAAAAAAA 2869

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estlom:*
5: em_estlpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_estl:*
11: gb_estl2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	766.6	26.7	947	11	B1156115 602903680
2	757	26.4	823	11	BG922983 602823860
3	752.2	26.2	788	11	BG862148 602795715
4	751.4	26.2	1036	11	B1156246 602903410
5	729.4	25.4	840	11	BG975502 602842871
6	718.8	25.1	731	11	BG920652 602826215
7	708.2	24.7	869	11	BG174177 602334170
8	702.6	24.5	889	11	BG175965 602337885
9	702.2	24.5	960	11	B1112360 602900049
10	700.2	24.4	717	11	BG976226 602846477
11	697	24.3	2634	12	AK018544 Mus muscu
12	694.4	24.2	719	11	B1108275 602902137

13	688.4	24.0	714	11	BG973911 602843546
14	679	23.7	736	11	B1155944 602904278
15	676.6	23.6	717	11	B1113023 602898184
16	675.6	23.5	808	11	BG914848 602813669
17	673	23.5	735	11	BG963048 602828060
18	671.2	23.4	688	11	B1157660 602920542
19	668	23.3	950	11	BG323650 602421932
20	663.2	23.1	1181	11	BG174568 602334411
21	657.6	22.9	744	11	B1219572 602936656
22	652.6	22.7	841	11	B141268 601789580
23	652	22.7	666	11	B157609 602922581
24	651.8	22.7	1186	11	BE794248 601591520
25	648.2	22.6	882	11	BE101658 601753581
26	646	22.5	797	10	A1324586 ms23n04.y
27	643.8	22.4	1073	11	BE913188 601668179
28	641.6	22.4	797	11	BG076355 H3158G04-
29	627	21.9	1137	11	BF135796 601782085
30	625.4	21.8	961	11	BF181575 601808876
31	623.6	21.7	911	11	BF141093 601786181
32	622	21.7	724	11	B1153867 602870869
33	616.6	21.5	672	11	BG146440 mab91f05-
34	614.4	21.4	616	11	BG864530 602798560
35	613.8	21.4	724	11	BE794948 601588709
36	613	21.4	736	10	AU132777 AU132777
37	607.6	21.2	902	11	BG756243 602713555
38	607.4	21.2	957	11	BF161870 601766802
39	603.8	21.0	1123	11	BF143134 601788026
40	602.6	21.0	676	10	BE374287 601227646
41	599.4	20.9	802	11	BF162704 601769359
42	597.2	20.8	772	11	BG914463 602813202
43	594.6	20.7	931	11	BF142884 601787952
44	591.6	20.6	656	10	BE284863 601097521
45	587.8	20.5	782	11	BF584946 602098830

ALIGNMENTS

RESULT 1
B1156115 947 bp mRNA
LOCUS 602901680P1 NIH.CGAP_Mam3 Mus musculus cDNA clone IMAGE:5033346 5',
DEFINITION B1156115
ACCESSION B1156115
VERSION B1156115.1 GI:14616116
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1. http://img.ncbi.nih.gov/Unpublished (1999)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1092 row: P column: 19
High quality sequence stop: 772.
location/Qualifiers
1. 947
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5033346"
/clone_lib="NIH.CGAP_Mam3"
/tissue="tumor, gross tissue"
/lab_host="DH10B"

FEATURES

source


```
QY 1970 CCAGTCTGTAGAGCCATFACACCAAGCAGACGCTGAACAACATGTCTATTGCTGAATCAT 2029
    |||||||
Db 241 CCAGTCTGTAGAGCCATFACACCAAGCAGACGCTGAACAACATGTCTATTGCTGAATCAT 300
QY 2030 CATGGCCTTAAGATCATGATGCGACCAACATCCGTGGTGTCCGCTGTCTACTCTA 2089
    |||||||
Db 301 CATGGCCTTAAGATCATGATGCGACCAACATCCGTGGTGTCTCTCACTGTCTACTCTA 360
QY 2090 CCCCGCATTTCCCAAGAGAGAGCATTTGGAAAGTACTGTAGGCCCGAGAGCCAGAGACA 2149
    |||||||
Db 361 CCCCGCATTTCCCAAGAGAGAGCATTTGGAAAGTACTGTAGGCCCGAGAGCCAGAGACA 420
QY 2150 CCCCCAAGCCGACCCAGTAGTGTCTCCCGCTACCTGAGACCAAGTGTATCTGTGTGAC 2209
    |||||||
Db 421 CCCCCAAGCCGACCCAGTAGTGTCTCCCGCTACCTGAGACCAAGTGTATCTGTGTGAC 480
QY 2210 ACCAAGCAGCTGACCAATACCAATTGACCTGCCGATGTCCCGGACATTTAGATTCTAT 2269
    |||||||
Db 481 ACCAAGCAGCTGACCAATACCAATTGACCTGCCGATGTCCCGGACATTTAGATTCTAT 540
QY 2270 GATGCACTTTGGAATTAACGCGTGAAGTGTCTGACCCCTGACGAGAGGCGAGTTGAGTC 2329
    |||||||
Db 541 GATGCACTTTGGAATTAACGCGTGAAGTGTCTGACCCCTGACGAGAGGCGAGTTGAGTC 600
QY 2330 GCTCAGCTTTGACATGATGATCTGACCTCGAGTGTCTACTCCCGCATGTGAGAGAGTGA 2389
    |||||||
Db 601 GCTCAGCTTTGACATGATGATCTGACCTCGAGTGTCTACTCCCGCATGTGAGAGAGTGA 660
QY 2390 AACCCAGACCTGACAGACGCTGACCTTGAGACACCTGCCCTGCTCCACCCCTTAACAGAC 2449
    |||||||
Db 661 AACCCAGACCTGACAGACGCTGACCTTGAGACACCTGCCCTGCTCCACCCCTTAACAGAC 719
QY 2450 CGAACCCCATATGCTGTGAAGACCTTAACCTTTGTGTGTCAGATTTTGTATTAATTC 2509
    |||||||
Db 720 -GAACCCATATGCTGTGAAGACCTTAACCTTTGTGTGTCAGATTTTGTATTAATTC 777
QY 2510 CTACTCTCTGCTATCTTTGGGCAATCTGGGCACTTTTAAAGA 2552
    |||||||
Db 778 CTACTCTCTGCTATC-TTGGGCAATCTGGGCTTTTAAATGA 819

RESULT 3
Bg862148 788 bp mRNA EST 29-MAY-2001
LOCUS 602795715F1 NIH_CCAP_Mam4 Mus musculus cDNA clone IMAGE:4916931 5',
DEFINITION mRNA sequence.
ACCESSION Bg862148
VERSION Bg862148.1 GI:14212686
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 788)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL NIH-MGC http://mgi.nci.nih.gov/
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LHAM0837 row: n column: 04
High quality sequence start: 2
High quality sequence stop: 781.
Location/Qualifiers
1. .788
/organism="Mus musculus"
```

```
/db_xref="taxon:10090"
/clone="IMAGE:4916931"
/clone_1lb="NIH_CCAP_Mam4"
/tissue_type="tumor, gross tissue"
/1ab_host="PH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: NCI;
Site_2: Sall; Cloned unidirectionally. Primer: Oligo dT.
Average insert 2.5 kb. Library constructed by Life
Technologies, catalog # 12018-016. Investigators providing
samples: Lothar Hennighausen/Priscilla Furth, NIH
Reference for Transgenic model: Li et al., Cell Growth and
Differentiation 7, 3-11 (1996). Note: this is a NCI-CCAP
library."
BASE COUNT 209 a 191 c 216 g 172 t
ORIGIN
```

```
Query Match 26.2%; Score 752.2; DB 11: Length 788;
Best Local Similarity 98.7%; Pred. No. 3,3e-162;
Matches 779; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
```

```
QY 1135 CTGAGTTGAATTATACAGCTTAAATTAAGTGTGATGAATAAGACTCTGGGATGTTG 1194
    |||||||
Db 1 CTGAGTTGAATTATACAGCTTAAATTAAGTGTGATGAATAAGACTCTGGGATGTTG 60
QY 1195 CTGCCCTCAGAGGGGTCTCGGAATTTAACATTTCTGGGCACGAACAAAAGTATGAGACA 1254
    |||||||
Db 61 CTGCCCTCAGAGGGGTCTCGGAATTTAACATTTCTGGGCACGAACAAAAGTATGAGACA 120
QY 1255 TGGAGAGCTTAACAACAGGCGCTCTGTGAGAGTTCAAGACCTGACTTATAGGAGC 1314
    |||||||
Db 121 TGGAGAGCTTAACAACAGGCGCTCTGTGAGAGTTCAAGACCTGACTTATAGGAGC 180
QY 1315 AGAGATGTGGGAATGGAGGCGCTGCCAATTTGTATGCTCTTGATGCTGACTGAGAGC 1374
    |||||||
Db 181 AGAGATGTGGGAATGGAGGCGCTGCCAATTTGTATGCTCTTGATGCTGACTGAGAGC 240
QY 1375 TGCACCTGATCAGCTTCGAGAGCTGAGGTGATGACCAAGGCTCAAGATTGACCTAGAGA 1434
    |||||||
Db 241 TGCACCTGATCAGCTTCGAGAGCTGAGGTGATGACCAAGGCTCAAGATTGACCTAGAGA 300
QY 1435 CCCAATTTTGGCCAGTGTGTGATCTCCCAACATCTGTCAATGCCAATGCTTGGGCTAT 1494
    |||||||
Db 301 CCCAATTTTGGCCAGTGTGTGATCTCCCAACATCTGTCAATGCCAATGCTTGGGCTAT 360
QY 1495 CAATTTTGTATACATCTGACCAATTAACCCCAAGAAATTTACTTCTGACTAAGC 1554
    |||||||
Db 361 CAATTTTGTATACATCTGACCAATTAACCCCAAGAAATTTACTTCTGACTAAGC 420
QY 1555 CGTACCTGGAGCAATGTGGCGGAGGTGCTCAGTCTCTGCTCCACCA 1614
    |||||||
Db 421 CGTACCTGGAGCAATGTGGCGGAGGTGCTCAGTCTCTGCTCCACCA 480
QY 1615 CCAATCTGAGCATGAGCAGCTGACCAACGCTGGCTTAAGCTCTTGAAGGCTG 1674
    |||||||
Db 481 CCAATCTGAGCATGAGCAGCTGACCAACGCTGGCTTAAGCTCTTGAAGGCTG 540
QY 1675 GTCATCTAGGCTGTACATGACATGGGCTAAATTCCTTAAGAAACATGGCTG 1734
    |||||||
Db 541 GTCATCTAGGCTGTACATGACATGGGCTAAATTCCTTAAGAAACATGGCTG 600
QY 1735 GCAATGATCTTCTCTGCTGTGGGTGTGATGACCAATATCATGACCTGTGAAAAGATATA 1794
    |||||||
Db 601 GCAATGATCTTCTCTGCTGTGGGTGTGATGACCAATATCATGACCTGTGAAAAGATATA 660
QY 1795 TCTTGGCCCTTTGGAAATGAAGGGTATCATGAGTGTTCATCAGCAAGAGCGGAGCGGG 1854
    |||||||
Db 661 TCTTGGCCCTTTGGAAATGAAGGGTATCATGAGTGTTCATCAGCAAGAGCGGAGCGGG 720
QY 1855 CCATCTTAAGCACAAAGCCCGGCGACCTTCTACTGCGCTTACGAGAGAGAGCAAGAA 1913
    |||||||
Db 721 CCATCTTAAGCACAAAGCCCGGCGACCTTCTACTGCGCTTACGAGAGAGAGCAAG 779
```

Oy	1914	GAAGACGGC	1922	
		1	111	
Db	780	AAGACGGC	788	
RESULT	4			
LOCUS	B1156246			
DEFINITION	B1156246	1036 bp	mRNA	EST
ACCESSION	602903410p1	NIH_CGAP_Mam3	Mus musculus	cdna clone IMAGE:5033019 5'
VERSION	B1156246			
KEYWORDS	B1156246.1	GI:14616247		
SOURCE	EST			
ORGANISM	house mouse.			
COMMENT	Mus musculus.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	1 (bases 1 to 1036)			
JOURNAL	NIH-MSGC http://msgc.nci.nih.gov/ .			
	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1993)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cga@bsi-remail.nih.gov			
	Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.			
	CDNA Library Preparation: Life Technologies, Inc.			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLM11092 row: c column: 04			
	High quality sequence start: 12			
	High quality sequence stop: 826.			
FEATURES	Location/Qualifiers			
source	1..1036			
	/organism="Mus musculus"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:5033019"			
	/clone_lib="NIH_CGAP_Mam3"			
	/tissue_type="tumor, gross tissue"			
	/lab_host="DH10B"			
	/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;			
	Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.			
	Average insert 2 kb. Library constructed by Life			
	Technologies, catalog #12017-018. Investigators providing			
	samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference			
	for transgenic model: Xu et al., Nature Genetics 22, 37-43			
	(1999). Note: This is a NCI/CGAP Library."			
BASE COUNT	238 a	287 c	271 g	240 t
ORIGIN				

Query Match	26.2%	Score 751.4	DB 11	Length 1036
Best Local Similarity	90.0%	Pred. No. 4,8e-162		
Matches 874	Conservative	0	Mismatches 86	Indels 11
			Gaps	6
QY	1725	AAACATGCTGGCAAGGCTCTCTCTCTGTGGCTGTGCTAGACAAATATATCATGACCTTGTG	1784	
DB	29	ACGCTGCCACCGACGCTGTTCTCTCTGTGGCTGTGACAAATATATCATGACCTTGTG	88	
QY	1785	AAAAATATATCTTGCCCTTTGGATGAAGGTGCATATGAGGTTTATCAGCAAGGAG	1844	
DB	89	AAAAATATATCTTGCCCTTTGGATGAAGGTGCATATATGAGGTTTATCAGCAAGGAG	148	
QY	1845	CGGAGACGGGCCATCTCTAAGCACAAAGCCCGGGACCTCTCTACGTGGCTTCAGCGAG	1904	
DB	149	CGGAGACGGGCCATCTCTAAGCACAAAGCCCGGGACCTCTCTACGTGGCTTCAGCGAG	208	
QY	1905	ACGAGCAAGAAGACGAGGGTCACCTTACATTGGGTGGAAAGACATCATGTGGCAAGACC	1964	
DB	209	ACGAGCAAGAAGACGAGGGTCACCTTACATTGGGTGGAAAGACATCATGTGGCAAGACC	268	
QY	1965	CAGATCCAGTCTGTAGAGCCATACACCAAGCAGCAGCTTAACACATGATCATTTGTGTA	2024	

[illegible]

RESULT	5				
LOCUS	BG975502				
DEFINITION	602842871f1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4978745 5'				
ACCESSION	BG975502	840 bp	mRNA	EST	12-JUN-2001
VERSION	BG975502.1				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; 1 (bases 1 to 840)				
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-r@mail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch Ph.D. cDNA Library Preparation: Life Technologies, Inc.				

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MCG clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL0975 row: m column: 18
Pin quality sequence stop: 815

BASE COUNT	220 a	211 c	226 g	183 f
ORIGIN				

Query Match	25.4%;	Score 729.4;	DB 11;	Length 840;
Best Local Similarity	95.9%;	Pred. No. 5.6e-157;		
Unshared	902;	Complementation	0;	Mismatches 26
				Indels 8;
				Coverage 5

QY	1802	CGTTTGGAATGAAGGGTACATCATCATCGCTTTATATCGCAAGGAGCGGAGCGGCATCT	1861
Db	2	CCCTTGGAATGAAGGGTACATCATCATCGCTTTATATCATGCAAGGAGCGGAGCGGCATCT	61
QY	1862	AAGCACAAGCCCCCGGCGACCTTCCTCTACTGCGCTTACAGCCGAGACGACGCAAGAAGAGAG	1921
Db	62	AAGCACAAGCCCCCGGCGACCTTCCTCTACTGCGCTTACAGCCGAGACGACGCAAGAAGAGAG	121
QY	1922	GGTCATCTTCACCTTGGGGTGAAGGAAGACATGATGGCAAGACCCGATCCAGTCTGTAGA	1981
Db	122	GGTCATCTTCACCTTGGGGTGAAGGAAGACATGATGGCAAGACCCGATCCAGTCTGTAGA	181
QY	1982	GCCATACACCAAGCAGCAGCTGAAACACATGTCATTTGCTGTGAATCATCATGCGCTATPA	2041
Db	182	GCCATACACCAAGCAGCAGCTGAAACACATGTCATTTGCTGTGAATCATCATGCGCTATPA	241
QY	2042	GATCATGTGATGGAGCAACATCCCTGGTGTCTGCATCTTGCTACCTCTACCCGACATTC	2101
Db	242	GATCATGTGATGGAGCAACATCTGGTGTCTGCATCTTGCTACCTCTACCCGACATTC	301
QY	2102	CAAGAGGAGGCGATTGGAAAGTACTGTTAGGCGCCGAGAGCCAGAGCAACCCGGAACCGA	2161
Db	302	CAAGAGGAGGCGATTGGAAAGTACTGTTAGGCGCCGAGAGCCAGAGCAACCCGGAACCGA	361
QY	2162	CCGAGTGTGTGGTGGCCCGGTACCTGAAACCAAGTTCACTGTGTGACACCAAGACCTG	2221
Db	362	CCGAGTGTGTGGTGGCCCGGTACCTGAAACCAAGTTCACTGTGTGACACCAAGACCTG	421
QY	2222	CAGCAATACCATTTGACCTCGCCGATGTCCCGCCGCACTTTAGATTTCATTGATGACGTTTGG	2281
Db	422	CAGCAATACCATTTGACCTCGCCGATGTCCCGCCGCACTTTAGATTTCATTGATGACGTTTGG	481
QY	2282	AAATTAAGGTGAAGTGGTGAACCTTCAGACAGAGAGGAGAGTTTGAAGTGGCTACGTTTGA	2341
Db	482	AAATTAAGGTGAAGTGGTGAACCTTCAGACAGAGAGGAGAGTTTGAAGTGGCTACGTTTGA	541
QY	2342	CATGATATCGAAGCTGGAGTGTGCTACCTCCGCCATGTGAGAGAGCTGAACAACGAGAAGCTG	2401
Db	542	CATGATATCGAAGCTGGAGTGTGCTACCTCCGCCATGTGAGAGAGCTGAACAACGAGAAGCTG	601
QY	2402	CAGAGAGTGAAGTGAACACCTGGCCC-GTGCTTCACCCCTAAGCAGCCGACACCCATA	2460

Db	602	CAGAAATTTGACTTGAGAACACCTGCCCGGTCCTCAACCCTAACGAGCGAA--CCCATATA	660
QY	2461	TCGCTGAAACTCCTAACATTTGTGGTCCAGATT--TTTTTTTAAATTTCCTACTTCTG	2518
Db	661	TCGCTGAAACTCCTAACATTTGTGGTCCAGATTATTTATTTCTCTACTTCTG	720
QY	2519	CTATCTTTGGGCAATCTGGGCACTTTTAAAAAGAGAAATGAGAGCTGCGGTGATAA	2578
Db	721	GTATCTTTGGGCAATCTGGGCACTTTTAAAAATGAGAAATGAGAGTGTGGGTGATTA	780
QY	2579	ACTGTTATGTAAGAGGAGACACCTCTGACAGCTGGGGATGGGGCTGAGAGCGAAGATG	2635
Db	781	--CTTATCTAAGAGGAGAC--TTGAGTGTGGGATGGGGGTGAAGCGAAGGGG	833

RESULT	6
BG920652	
LOCUS	BC920652 731 bp mRNA
DEFINITION	G02820215.1 NCI_CGAP_Mam6 Mus musculus cDNA IMAGE:4954946 5'
ACCESSION	mRNA sequence.
VERSION	BC920652
KEYWORDS	BC920652.1 GI:14301128
SOURCE	EST.
ORGANISM	house mouse. Mus musculus

BASE COUNT	191 a	183 c	198 g	158 t	1 others
ORIGIN					

Query Match	25.1%;	Score 718.8;	DB 11;	Length 731;
Best Local Similarity	98.9%;	Pred. No. 1.6e-154;		
Matches 723;	Conservative	0;	Mismatches 8;	Indels 0;
				Gaps 0;

QY	899	QY	899
CCGCGCAACAAATTAAAGAAACGTGAGAGAGCTCGAGCGAAAGTGTCTACAAAGCGGACCC	958	CCGCGCAACAAATTAAAGAAACGTGAGAGAGCTCGAGCGAAAGTGTCTACAAAGCGGACCC	958
CCGCGCAACAAATTAAAGAAACGTGAGAGAGCTCGAGCGAAAGTGTCTACAAAGCGGACCC	60	CCGCGCAACAAATTAAAGAAACGTGAGAGAGCTCGAGCGAAAGTGTCTACAAAGCGGACCC	60
QY <td>959</td> <td>QY <td>959</td> </td>	959	QY <td>959</td>	959
TATCGTGCTGACGACCCGCCATGCTCGAGAGAGAGATCGTGGACGTGTTTCAGAACTTAAT	1018	TATCGTGCTGACGACCCGCCATGCTCGAGAGAGAGATCGTGGACGTGTTTCAGAACTTAAT	1018
TATCGTGCTGACGACCCGCCATGCTCGAGAGAGAGATCGTGGACGTGTTTCAGAACTTAAT	120	TATCGTGCTGACGACCCGCCATGCTCGAGAGAGAGATCGTGGACGTGTTTCAGAACTTAAT	120

Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db 1019	GAAGATGCTTCTGTTGTGAGCGGACAGCCCTGCATGCCCCATGCACCCGAGCCGCTT	1078					
Db 121	GAAGATGCTTCTGTTGTGAGCGGACAGCCCTGCATGCCCCATGCACCCGAGCCGCTT	180					
Oy 1079	AGTCATCAAGACCTGTGTGCCAGTTTACCACCAAAATCAGTTTGCCTGTCATATTTCTCTGA	1138					
Db 181	AGTCATCAAGACCTGTGTGCCAGTTTACCACCAAAATCAGTTTGCCTGTCATATTTCTCTGA	240					
Oy 1139	GTTCGAATTAATCAGCTTAAATTAAGTGTGCATTTGATTAAGACTCTGGGGATGTTTCTGC	1198					
Db 241	GTTCGAATTAATCAGCTTAAATTAAGTGTGCATTTGATTAAGACTCTGGGGATGTTTCTGC	300					
Oy 1199	CCTGAGAGGCTCTGGGAATTTAACTATTCTGGGACACGAACACAAAGATGATGATCATGA	1258					
Db 301	CCTGAGAGGCTCTGGGAATTTAACTATTCTGGGACACGAACACAAAGATGATGATCATGA	360					
Oy 1259	GGAGCTTAACAACGCGCAGCTGTCTGTGCAGAGTTCAAGCACTGACCTTTAGGGACAGAG	1318					
Db 361	GGAGCTTAACAACGCGCAGCTGTCTGTGCAGAGTTCAAGCACTGACCTTTAGGGACAGAG	420					
Oy 1319	ATGTGGAATGAGAGCGCTGCCAATTTGATGATCCCTGATGATGCTGAGCTGAGAGCTTGA	1378					
Db 421	ATGTGGAATGAGAGCGCTGCCAATTTGATGATCCCTGATGATGCTGAGCTGAGAGCTTGA	480					
Oy 1379	CCTGATCACTTTGAGACTGAGGTGTACCACCAAGGCTTCAAGATTGACTTGAACCCA	1438					
Db 481	CCTGATCACTTTGAGACTGAGGTGTACCACCAAGGCTTCAAGATTGACTTGAACCCA	540					
Oy 1439	CTCCCTGCGAGTTGGTGATGATCCCAATCATGTGATGATGCCAATGCTTGGGACTATCAAT	1498					
Db 541	CTCCCTGCGAGTTGGTGATGATCCCAATCATGTGATGATGCCAATGCTTGGGACTATCAAT	600					
Oy 1499	CCTGTGTATTAACATGCTGACCAATTAACCCCAAGAACCTGTAATTAAGCCGCC	1558					
Db 601	CCTGTGTATTAACATGCTGACCAATTAACCCCAAGAACCTGTAATTAAGCCGCC	660					
Oy 1559	AATTGAACTGTGGAGCAAGTGGCCGAGTGTCTAGTGGCATTTCTCTCACCACCA	1618					
Db 661	AATTGAACTGTGGAGCAAGTGGCCGAGTGTCTAGTGGCATTTCTCTCACCACCA	720					
Oy 1619	CGGAGCGCTGA	1629					
Db 721	CGGAGCGCTGA	731					
RESULT 7							
LOCUS	BG174177	869 bp	mRNA	EST	06-FEB-2001		
DEFINITION	602334170F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457248	5'					
ACCESSION	BG174177						
VERSION	BG174177.1	GI:12680880					
KEYWORDS	EST						
SOURCE	house mouse.						
ORGANISM	Mus musculus						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
TITLE	1 (bases 1 to 869)						
JOURNAL	NIH-MGC http://mgi.nci.nih.gov/						
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)						
	Unpublished (1999)						
	Contact: Robert Strausberg, Ph.D.						
	Email: cgabs-remail.nih.gov						
	Tissue Procurement: Gilbert Smith, Ph.D.						
	cDNA Library Preparation: Life Technologies, Inc.						
	DNA Sequencing by: Incyte Genomics, Inc.						
	Clone distribution: MGC clone distribution information can be						
	found through the I.M.A.G.E. Consortium/LLNL at:						
	http://image.llnl.gov						
	Plate: LLM10253	row: 1	column: 17				
	High quality sequence stop: 720.						

FEATURES					
	Source	Location/Qualifiers			
		1..869	/organism="Mus musculus"		
			/strain="FVB/N"		
			/db_xref="taxon:10090"		
			/clone_image:4457248"		
			/clone_lib="NCI CGAP_Mam1"		
			/tissue_type="tumor, biopsy sample"		
			/dev_stage="3 months, virgin"		
			/lab_host="IMH0B"		
			/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"		
	BASE COUNT	245 a	220 c	253 g	151 t
	ORIGIN				
	Query Match	24.7%	Score 708.2:	DB 11:	Length 869;
	Best Local Similarity	97.1%:	Pred. No. 4.1e-152:		
	Matches 764:	Conservative	0:	Mismatches 18:	Indels 5: Gaps 4:
OY	407	TCGCCTCCTCCAGACGGCAGCGGCAACCAGCAAGGGGGGCCAAGGCCAACCCCAAC	466		
Db	1	TCGCCCTCCCTCCAGAGGGCGCCAGCGCACGCCAGCAAGGGGGGCCAAGCCCAAC	60		
OY	467	AGCCGCGGTAGTGACAGAGAACAACACACATGTTGGAGCAGCATCTTCACAGATGTCGGAA	526		
Db	61	AACCCCGGTAGTGACAGAGAACAACACACATGTTGGAGCAGCATCTTCAGAGATGTCGGAA	120		
OY	527	GCGAGTGCAGGACTTGAAGACAGAAAATGAGGTGGTGAGAACCTCCAGAGACACTTTGA	586		
Db	121	GCGAGTGC AGGACTT GAAGACAGAAAATGAGGTGGTGAGAACCTCCAGAGACACTTTGA	180		
OY	587	TTCGAACTCAAAAACCCCTTAAGAGCCAAGAGACATCGAGATCTGGAATGATAACACCA	646		
Db	181	TTTCGAACTCAAAAACCCCTTAAGAGCCAAGAGACATCGAGATCTGGAATGATAACACCA	240		
OY	647	GTCTGTGACACAGACAGAAGATGCACAGCTGGAGACAGATGCTCACAGCCCTTGACCAGAT	706		
Db	241	GTCTGTGACACAGACAGAAGATGCACAGCTGGAGACAGATGCTCACAGCCCTTGACCAGAT	300		
OY	707	GCGGGAACAACATTGTGAGTAGTGAGCTGGCGGGGGCTCTTGTAGCAATGGAGTACGTCCAGAA	766		
Db	301	GCGGGAACAACATTGTGAGTAGTGAGCTGGCGGGGGCTCTTGTAGCAATGGAGTACGTCCAGAA	360		
OY	767	GACACGTAGTATGAAGACTGGCTGTACTGGAAGAAGCGCGCCAGAGATCGCTGCATCGG	826		
Db	361	GACACGTAGTATGAAGACTGGCTGTACTGGAAGAAGCGCGCCAGAGATCGCTGCATCGG	420		
OY	827	AGGCCCTCCCACACATCTGCCCTGGACCGCTCTGTGAAAACTGGATTACTTATTTAGCAGAAATC	886		
Db	421	AGGCCCTCCCACACATCTGCCCTGGACCGCTCTGTGAAAACTGGATTACTTATTTAGCAGAAATC	480		
OY	887	TCGACTTCAGACCCCGCCCAAAATTTAAGAACTGGAGAGCTGTGCAGCAAAAGTGTCTTA	946		
Db	481	TCGACTTCAGACCCCGCCCAAAATTTAAGAACTGGAGAGCTGTGCAGCAAAAGTGTCTTA	540		
OY	947	CAGAGCGGACCCCTATGTGTGCACACCGCGCCACTGCTGAGAGAGAGANTGTGGAGACTGTT	1006		
Db	541	CAAGGCGGACCCCTATGTGTGCACACCGCGCCACTGCTGAGAGAGAGANTGTGGAGACTGTT	600		
OY	1007	CAGAAACTTAATGAAGAGTCCCTGCTGTGTGAAGCGGCAAGCCCTGATGCCATGCATGACCC	1066		
Db	601	CAGAAACTTAATGAAGAGTCCCTGCTGTGTGAAGCGGCAAGCCCTGATGCCATGCATGACCC	660		
OY	1067	GGACCGGCGCCTTATGATATGAAGACT -GGTGTCCAGTTTACG -ACGAAATCGAGTGTGCTG	1124		
Db	661	GGACCGGCGCCTTATGATATGAAGACTGGGTGTCTCAAGTATTACCAACGAAATCGAGGGTGGT	720		
OY	1125	GTCGAATTTCTCGA -GTCGAATTATCAGCTT - AAAATTTAAGGTGTGATTCATTAAGAC	1181		
Db	721	GTCGAATTTCTCGAGTTGAATATATCAGCTT TTTAAATTTACAGGGTGGCCATATCAAGAC	780		

```
OY      1182 TCTGGGG 1188
|||||
Db      781 TCTGGGG 787

RESULT  8
Bg175965      889 bp      mRNA      EST      06-FEB-2001
LOCUS      602337885F1 NCI_CGAP_Mam1 Mus musculus cdna clone IMAGE:4461020 5',
DEFINITION      mRNA sequence.
ACCESSION      Bg175965
VERSION      Bg175965.1 GI:12682668
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  1 (bases 1 to 889)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgaps-r@mail.nih.gov
              Tissue Procurement: Gilbert Smith, Ph.D.
              CDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: LLAM10263 row: i column: 21
              High quality sequence stop: 714.
              Location/Qualifiers
                1..889
                /organism="Mus musculus"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:4461020"
                /clone_1lb="NCI_CGAP_Mam1"
                /tissue_type="tumor, biopsy sample"
                /dev_stage="3 months, virgin"
                /lab_host="DH10B"
                /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
                Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                Library constructed by Life Technologies. Investigator
                providing samples: Gilbert Smith, NIH"

BASE COUNT      221 a      232 c      242 g      194 t
ORIGIN

Query Match      24.5%: Score 702.6; DB 11: Length 889;
Best Local Similarity 96.9%: Pred. No. 7, Be-151;
Matches 770; Conservative 0; Mismatches 19; Indels 6; Gaps 5;

OY      1177 GCAAGAAACATGCTGCTGAGGCGCTTCTCTTGTGCTGTGCTAGACAATATCATCG 1776
|||||
Db      1      GCAAGAAACATGCTGCTGAGGCGCTTCTCTTGTGCTGTGCTAGACAATATCATCG 60

OY      1777 ACCTTGAAAAAGATATCTCTGGCCCTTGGAAATGAAGGTACATGCTGGTTTCATCA 1836
|||||
Db      61      ACCTTGAAAAAGATATCTCTGGCCCTTGGAAATGAAGGTACATGCTGGTTTCATCA 120

OY      1837 GCAAGAGCGGAGCGGCCATCTTAAGCACAAAGCCCCCGGACCTTCTTACTGCGCT 1896
|||||
Db      121      GCAAGAGCGGAGCGGCCATCTTAAGCACAAAGCCCCCGGACCTTCTTACTGCGCT 180

OY      1897 TCACGAGACAGACGAAAGAGAGGGGTCACTTTCACCTTGGTGGAAAGACATCACTG 1956
|||||
Db      181      TCACGAGAGACGAAAGAGAGGGGTCACTTTCACCTTGGTGGAAAGACATCACTG 240

OY      1957 GCAAGACCCAGATCCAGTCTGTAGAGCCATACCAAGACACAGCTGAACAATGTCTAT 2016
|||||
```

```
Db      241      GCAAGACCCAGATCCAGTCTGTAGAGCCATACACCAAGACGACGCTGACACATGTCTAT 300

OY      2017 TTGCTGAATATCATGATGCGCTATATAAGATCATGATGCGACCAACATCCTGGTGTCTCAC 2076
|||||
Db      301      TTGCTGAATATCATGATGCGCTATATAAGATCATGATGCGACCAACATCCTGGTGTCTCAC 360

OY      2077 TTGCTACCTCTACCCGACATTCGCCAAGAGAGGAGGATTTGGAATACCTAGTGAAGGCCCG 2136
|||||
Db      361      TTGCTACCTCTACCCGACATTCGCCAAGAGAGGAGGATTTGGAATACCTAGTGAAGGCCCG 420

OY      2137 AGAGCCAGAGACACCCCGAAGCCGACCCAGTAGTGTGCCCCGTACCTGAAGACCAAGT 2196
|||||
Db      421      AGAGCCGAGAGACCCCGAAGCCGACCCAGTAGTGTGCCCCGTACCTGAAGACCAAGT 480

OY      2197 TCATCTGTGTACACCAACGACCTGACGAATATACCTTGTACCTGCGATGTCCCGCGCA 2256
|||||
Db      481      TCATCTGTGTGTACACCAACGACCTGACGAATATACCTTGTACCTGCGATGTCCCGCGCA 540

OY      2257 CTTAGATTCATTTGATGACGATTTGGAATTAACGGGTGAAGGCTGAGGCCCTCAGCAGAGAG 2316
|||||
Db      541      CTTAGATTCATTTGATGACGATTTGGAATTAACGGGTGAAGGCTGAGGCCCTCAGCAGAGAG 600

OY      2317 GCGAGTTTGAGTCGCTGACGTTTGACATGATGTC-TGACCTCGAGATG-TCATACCTCCC 2374
|||||
Db      601      GCGAGTTTGAGTCGCTGACGTTTGACATGATGTC-TGACCTCGAGATG-TCATACCTCCC 660

OY      2375 CATGTGAGAGCTTAACCAACGACCTGCAGAGAGCTGACTTGAACACCTTCCCGCTGCT 2434
|||||
Db      661      CATGTGAGAGCTTAACCAACGACCTGCAGAGAGCTGACTTGAACAACTG-CCCGGCT 718

OY      2435 CCACCCCTTAAGCAGCCGACACCCATATGCTGTGAAGCTCAATCTTGTGCTGACATTT 2494
|||||
Db      719      CCACCCCTTAAGCAGCCGACCAACCATATGCTGTGAAGCTCAATCTTGTGCTGACATTT 776

OY      2495 TTTTGTTTTAATTC 2509
|||||
Db      777      TTTAATTCCTACTTC 791

RESULT  9
Bg112360      960 bp      mRNA      EST      26-JUN-2001
LOCUS      602900049F1 NCI_CGAP_Mam5 Mus musculus cdna clone IMAGE:5029548 5',
DEFINITION      mRNA sequence.
ACCESSION      Bg112360
VERSION      Bg112360.1 GI:14563261
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  1 (bases 1 to 960)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgaps-r@mail.nih.gov
              Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
              CDNA Library Preparation: Life Technologies, Inc.
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: LLAM1083 row: b column: 13
              High quality sequence stop: 789.
              Location/Qualifiers
                1..960
                /organism="Mus musculus"
                /strain="C57/B6"
                /db_xref="taxon:10090"
                /clone="IMAGE:5029548"
                /clone_1lb="NCI_CGAP_Mam5"
```


/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site_1: Sal1;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lochar Hennighausen/Kodin Humphreys,
NIH"

BASE COUNT 247 a 258 c 260 g 195 t
ORIGIN

Query Match 24.5%; Score 702.2; DB 11; Length 960;
Best Local Similarity 94.3%; Pred. No. 9.5e-151;
Matches 784; Conservative 0; Mismatches 38; Indels 9; Gaps 5;

1436 CCACCTCTTCCAGTGTGTGATCTCCACATCTCTCAGATCCCAATGCTTGGCATC 1495
|||||
4 CCACCTCTTCCAGTGTGTGATCTCTCCACATCTCTCAGATCCCAATGCTTGGCATC 63
1496 AATCTGTGTATACATGCTGACCAATAACCCCAAGACGTGAATCTTCACTAAGCC 1555
|||||
64 AATCTGTGTATACATGCTGACCAATAACCCCAAGACGTGAATCTTCACTAAGCC 123
1556 GCCAATTGGAACTGGGACCAAGTGGCCGAGTGTCTCAGCTGGCAGTTCTCTGACACAC 1615
|||||
124 GCCAATTGGAACTGGGACCAAGTGGCCGAGTGTCTCAGCTGGCAGTTCTCTGACACAC 183
1616 CAAGCGAGGCTGAGCATCGACAGCTGACACGCTGGCTGAGAAGCTCTAGGGCTGG 1675
|||||
184 CAAGCGAGGCTGAGCATCGACAGCTGACACGCTGGCTGAGAAGCTCTAGGGCTGG 243
1676 TGTGACTACTCAGAGGTGTAGATCAGATGGCTAAATTTCTGCAAGAAAACATGGCTGG 1735
|||||
244 TGTGAATCTCAGAGGTGTAGATCAGATGGCTAAATTTCTGCAAGAAAACATGGCTGG 303
1736 CAAGGCGTTCTCTTGGGTGGCTGCTAGACAAATATCATGACCTTGTGAAAAGTATAT 1795
|||||
304 CAAGGCGTTCTCTTGGGTGGCTGCTAGACAAATATCATGACCTTGTGAAAAGTATAT 363
1796 CTGGGCGTTTGGAAATGAAAGGGTACATCATGGTTTCATCAGCAAGAGCGGAGCGGC 1855
364 CTGGGCGTTTGGAAATGAAAGGGTACATCATGGTTTCATCAGCAAGAGCGGAGCGGC 423
1856 CATCTTAAGACAAAGCCCGGCGACCTTCTACTGCGCTTACGCGAGACAGCAAGAA 1915
|||||
424 CATCTTAAGACAAAGCCCGGCGACCTTCTACTGCGCTTACGCGAGACAGCAAGAA 483
1916 AGGAGGGTCACTTTCACCTTGGGTGGAAGAGCATAGTGGCAAGACCCAGATCCAGTC 1975
|||||
484 AGGAGGGTCACTTTCACCTTGGGTGGAAGAGCATAGTGGCAAGACCCAGATCCAGTC 543
1976 TGTAGAGCCATACACCAAGAGCAGCTGAACACATGTCTCATTTGCTAAATCATATGGG 2035
544 TGTAGAGCCATACACCAAGAGCAGCTGAACACATGTCTCATTTGCTAAATCATATGGG 603
2036 CTATAAATCATGATGATCGACCAACATCGTGTGTCT-CCACTGTCTACTCTTACCCCG 2094
604 CTATAAATCATGATGATCGACCAACATCGTGTGTCTCTCCACTGTCTTACTCTTACCCCG 663
2095 ACATTTCCCAAGAGAGGAGATTTT--GGAAAGTACTGTA--GCCCGAGAGCCAGAGATCACC 2151
664 ACATTTCCCAAGAGAGGAGGATTTGGAAAAGTCTCTGTAAGGGCCGAGAGCAGAGACACC 723
2152 CCGAAGCCGAGCCAGTGTGTGCTGCCCGTACTGGAAGACCA--GTTTCATCTGTGTGAC 2209
724 GCGAAGCCGAGCCAGTGTGTGCTGCCCGTACTGGAAGACCAAGATTCACTTGTGGTGAC 783
2210 ACCAAGAGCC--TGCAGCAATACGATTAAGTGGCGATGTGCCCCCGCAC 2257
784 ACCAAGAGCCGTGAGGCAATACCCATTGACCTGTCCGATGTGTCCCGCCCGC 834

RESULT 10
BG976226 717 bp mRNA EST 12-JUN-2001
LOCUS 602846477f1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4977115 5',
DEFINITION mRNA sequence.
ACCESSION BG976226
VERSION BG976226.1 GI:14363863
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 717)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM10971 row: 1 column: 20
High quality sequence stop: 699.

FEATURES
source
1..717
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_1lb="IMAGE:4977115"
/clone_1lb="NCI_CGAP_Mam6"
/sex="Female, virgin"
/tissue_type="inflicting ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site_1: Sal1;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 197 a 175 c 187 g 158 t
ORIGIN

Query Match 24.4%; Score 700.2; DB 11; Length 717;
Best Local Similarity 99.4%; Pred. No. 2.9e-150;
Matches 713; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

857 GGAAGCTGATTAATCTTATAGAGAAATCTCACTTCAAGCCGCGCAAAATTAAGAA 916
1 GGAAGCTGATTAATCTTATAGAGAAATCTCACTTCAAGCCGCGCAAAATTAAGAA 60
917 ACTGAGAGAGCTGACGAGAAAGTGTCTTCAAGGCGACCTTATCTCATCAAGAGTGTGT 976
61 ACTGAGAGAGCTGACGAGAAAGTGTCTTCAAGGCGACCTTATCTCATCAAGAGTGTGT 120
977 CATCTTGAGAGAGAGATCGTGAAGCTGTTGAGAACTTAATGAAGAGTGTCTTGTGT 1036
121 CATCTTGAGAGAGAGATCGTGAAGCTGTTGAGAACTTAATGAAGAGTGTCTTGTGT 180
1037 GAGAGCGGACCGCTGCAATGCCATGACACCGGACCGGCGCTTATCTCATCAAGAGTGTGT 1096
181 GGAGCGGACCGCTGCAATGCCATGACACCGGACCGGCGCTTATCTCATCAAGAGTGTGT 240
1097 CCAGTTTACCAAGAAAGTGTGCTGTGCTCAATTTCTGAGTGTGAATATACGCTTAA 1156
241 CCAATTTTACCAAGAAAGTGTGCTGTGCTCAATTTCTGAGTGTGAATATACGCTTAA 300
1157 AATTAAAGTGTGATTAAGAACTCTGGGAGTGTGCTGCTGCTCAGAGGGTCTCGGAA 1216
301 AATTAAAGTGTGATTAAGAACTCTGGGAGTGTGCTGCTGCTCAGAGGGTCTCGGAA 360

QY 1217 ATTTAATTCTGGGACGAGAACAAAGTATGACATGAGAGCTCTAACACGCGAG 1276
|||||
Db 361 ATTTAATTCTGGGACGAGAACAAAGTATGACATGAGAGCTCTAACACGCGAG 420
QY 1277 CCTGTGCGAGAGTTCAAGACCTGAGCCCTTGGAGAGAGATGTGGAAATGGAGCGG 1336
|||||
Db 421 CCTGTGCGAGAGTTCAAGACCTGAGCCCTTGGAGAGAGATGTGGAAATGGAGCGG 480
QY 1337 TGCCAAATTCGATCCCTCTTGTGATGCTGAGAGAGCTGACACCTGATCACCCTTGAGAC 1396
|||||
Db 481 TGCCAAATTCGATCCCTCTTGTGATGCTGAGAGAGCTGACACCTGATCACCCTTGAGAC 540
QY 1397 TGAGGTGACCAAGAGCGCTCAAGATTGACCTAGAGACACCACCTCTTGCCAGTTGTGCT 1456
|||||
Db 541 TGAGGTGACCAAGAGCGCTCAAGATTGACCTAGAGACACCACCTCTTGCCAGTTGTGCT 600
QY 1457 GATGTCCACATCTGTGATGATGAGAGCAATGCTGGGACATCAA-TCCGTGTGATTAACATGC 1515
|||||
Db 601 GATGTCCACATCTGTGATGATGAGAGCAATGCTGGGACATCAA-TCCGTGTGATTAACATGC 660
QY 1516 TGACCAATTAACCCAGAGAACTGTAACCTTCTTCACTAACCCGCAATTGCAACTGGG 1572
|||||
Db 661 TGACCAATTAACCCAGAGAACTGTAACCTTCTTCACTAACCCGCAATTGCAACTGGAG 717
RESULT 11
AK018544 2634 bp mRNA HTC 05-JUL-2001
LOCUS Mus musculus adult male colon cDNA, RIKEN full-length enriched
DEFINITION library, clone:9030611B09, full insert sequence.
ACCESSION AK018544
VERSION AK018544.1 GI:12858297
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male colon cDNA to mRNA,
clone:9030611B09.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 2634)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (bases 1 to 2634)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (bases 1 to 2634)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,K., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwa,K.,
Yamabe,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Okazaki,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
JOURNAL sequencing pipeline with 384 multicapillary sequencer
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4 (bases 1 to 2634)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 2634)

AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
Imotani,K., Ishii,Y., Itoh,M., Iwama,M., Kato,H., Kawai,J.,
Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
Sasaki,D., Shibata,K., Shibata,Y., Shihagawa,A., Shitaki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp,
URL:htp://genome-gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT Please visit our web site (<http://genome-gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
15' GAGAGAGAGAGAGATTCGAGATTAATAATGATCCCGCCCGCC 3'. cDNA was
prepared by using triethanolamine-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 10.0 and subtraction to
Rot = 185.2. Second strand cDNA was prepared with the primer
adapter of sequence 15'
GAGAGAGAGATTCGAGATTAATAATGATCCCGCCCGCC 3'. cDNA was cleaved
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
bulk excision from Lambda F1C 1. Cloning sites, 5' end: SalI; 3'
end: BamHI. Host: DH10B.
FEATURES
Source
Location/Qualifiers
1..2634
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGI:103063"
/db_xref="MGI:1907734"
/clone="9030611B09"
/sex="male"
/tissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
302..2551
/note="putative"
/protein_id="BAB31265.1"
/db_xref="GI:12858298"
/translation="MSQMFLEOGLDSEKFLKQVLYLDSPFMEIRQYLAQMLEKQME
HAADVSEFATREHDLISOLDQYSRSLLENFLLOINIRKSKRNLDNOEDPVQMS
MIYINCLKEERKILENQRQAEQGNIONTVLMDKRELDKVRNKKDDVCMCEDEL
KTLLELDEYDFCKTSQNRGEANGAKSDQKQELLHLMFIMLNKKREI1HKR
ELNLSIELTONTVLINDELVEKRRQSGACIGPNCALDLOJQSFVTAETLLOD1ROO
LKLELELQKFTPEPDIYTKNKKVLSRTEFLFQOLIOSSVVEVROCMPTHPORPLV
LKTGVOTVVKRLRLKLOELNYN1KAVSPDKQVNEKNTYKGFKNILGTHKRVNM
EESTNGSLAEFRHLOLKQKNGKRNTEPRLIYTERLSLSPETQICQGLVLDLET
TSLPVPVYISNVSQPSQMAS1LRYNMLYTERPRINSFLNPLPCAMWSQSLSEVLS
VTKRGLNADQLSM1GERCLUGPNAGPDL1PWTBRCNKINIDNKSEFWPW1DT1LELK
KHLCLVANDGCIWGFISKERBAL1KQOQDPTFLRPSSESRGAI1FTWVERSONG
EPDFHVAEPYTKELSAVTPPD1IRNRYVMAEEN1PNNPLKYL1YPNIDKHALGKYS
RPEKAPPEMELDPRKRGYIKTELISVSEVHPSRLQTTDNLPLPSPPEPDM5R1V
EPDSM5STV"

BASE COUNT 729 a 658 c 658 g 589 t
ORIGIN
Query Match 24.3%; Score 697; DB 12; Length 2634;
Best Local Similarity 60.4%; Pred. No. 1.2e-149;

Matches 1248: Conservative 0: Mismatches 795: Indels 24: Gaps 5:				
OY	66	AGATGGCTCTACTGGAACAGCTGCGACAGCTGGACACAGCTACTGTAAGCAGCTGCAC	125	
Db	299	AGGATGTACAGAGTGGTTCAGAGCTTCCAGAGCTTCCAGAGTTCCTGGGAGGTCAC	358	
OY	126	CAGCTTACAGGACAGCTTCCCATGGAGCTGGGAGTCTCGCACCCTTGGATTGAG	185	
Db	359	CAGCTTACAGTACAGCTTCCCATGGAGTACAGAGTACTGGGCGCCGCTGGAG	418	
OY	186	AGCAGAGCTGGGCATATGAGCCAGCAAGAGTACATGCGACGTTGGTTCATAT	245	
Db	419	AAGCAAGACTGGAGAGCAGCTGCTATGATGTCTGTTCCGACATCCCTTCATGAC	478	
OY	246	CTCTTGGGTAAATTGACCAAGCAATATAGCCGATCTCTGCAAGAGTCCATGTCCTAT	305	
Db	479	CTCTCTTCAAGCTGGAGCGCAGTACAGCCGCTTTCTGTGGAATTAATTTCTTGTG	538	
OY	306	CAGCACAACCTTGAAGATCAAGCAGTTCCTGAGAGCAGTATCTTGAAGCCAAATG	365	
Db	539	CAGCACAACATACGGAAGAACAGCCATATCTCCAGATTAATCTCCAGAGATCCGCTA	598	
OY	366	GAATTTGCCCGATCGTGCCCGATGCTGTGGAGAGTCTGCTCTCCAGACGGCA	425	
Db	599	CAGATGTGATGATCATCTACACTGTCTGAAGGAGAAAGAAATTTGAAATATGCC	658	
OY	426	GCCAGCGCAGCCAGCAAGGGGGCCAGCCACCCACCCAGCCGCGCTAGTGACAGAG	485	
Db	659	CAAGATTTTAATCA-----GGCCAGAGGGAATTTTCAGAACACTGTGATTTGAT	712	
OY	486	AAGCAGAGATGTTGGAGCAGCATCTTCAGATGTCCGGAAGCAGTGACAGATCTAGAA	545	
Db	713	AAMCAGAGAGAGCTGGACAGTAAAGTCAAGAAATGTGAAGATCAAGTCAATGTGATAG	772	
OY	546	CAGAAATGAAGTGTGTGAGAACCTCCAGAGCAGCTTTGATTTCAACTACAAACCTTC	605	
Db	773	CAGAAATCAAGACCTTAGAAGAAATACAGATGATATGACTTTAAATGCAAAACCTCT	832	
OY	606	AAGAGCCAGAGACATGCGAGATCTGAATGGAACAAACAGCTGTGACAGCAGAG	665	
Db	833	CAGAACGAGAG-----GTGAAGCCAAATGTGTGGGGAAGCCGACCAAAACAGGAA	886	
OY	666	ATGAGAGAGCTGGAACAGATGCTCACAGCCCTGAGACGATGCGAGAGCATTTGTAGT	725	
Db	887	CAGCTGCTGCTCCACAGATGTTTATGCTTGTACATTAAGAGAAAGAGATTAATTCAC	946	
OY	726	GAGCTGGCGGGCTCTTGTCAAGCAATGATGCTGCAAGAACACTGACTGATGAGAG	785	
Db	947	AAATTCAGTGAGTTGCTGAATTCATCGAGCTCACTCAGAACACTGTGATTAATGACAG	1006	
OY	786	CTGGCTACGTGAGAGGGGCGGAGATGGCTGATCGGAGGCGCTGCCAACATCTGC	845	
Db	1007	CTCTGGAGTGAAGCGAGGCGAGCGAGCCCTGCTGATGGGGGAGCCGCCAACGCTGCG	1066	
OY	846	CTGAGCGTCTGGAATACTGATTAATTCATTAGCAGAAATCTCAATTCAGACCCGCCAA	905	
Db	1067	CTGATCAGCTGCAAGAGTGTTCACCATTTGTGAGAGACCTTCAGACAGATCTCGTAC	1126	
OY	906	CAAAATTAAGAACTGGAGAGCTGCGAGAAAGTGTCTTACAGAGGGGAGCTATCTG	965	
Db	1127	CAGTTTAAAAAGCTGAGAGATTTGACAGAAATTCACCTATGAGCCCGACCTTAATTACA	1186	
OY	966	CAGAGCGGCCCTATGCTGGAGAGAGATCGTGAGCTGTTGAGAACTTAATGAGAGT	1025	
Db	1187	AAAAACAAGCAGGTGTTTTCAGATGAAACTTCTCTCTTTCAGACAGCTCATTTAGAGC	1246	
OY	1026	GCCTTGTGTGAGCGGAGCCCTGATGCCATGACACCCGAGCCGCGCTTACTATC	1085	
Db	1247	TCCTTGTGTGAGAGAGAGCGGCTGATGCCCTCACTCAGCCGACAGCCCGCTGTG	1306	
OY	1086	AAGACTGTGTCCAGTTACCAAGAAAGTCAAGTTGTGTGTAATTTCTGAGATTGAAT	1145	
Db	1307	AAGACTGGGTACAGTTCACTGTCAAGCTGAGACTGTGTGTGAATTTGCAAGAGTTGAA	1366	

OY	1146	TATCAGCTTAAATTAAGTGTGATGATTAAGACTCTGGGATGTTGTCCCTCAGA	1205	
Db	1367	TATACCTTGAAGTGAAGTCTCATTTGACAAAGATGTGAACAGAAACAGTTAA	1426	
OY	1206	GGGTCTCGAAATTTATCATTTCTGGGACAGACACAAAGTATGATGACATGGAGACT	1265	
Db	1427	GGATTTGGAAATTCACATCTTGGGTACGACACCAAAAGTATGAGCATGGAAGATTC	1486	
OY	1266	AACAGCGCCGCTGTGAGAGTTCAAGCACTGACCTTACGGAGCAGATGTGGC	1325	
Db	1487	ACCAAGGAAGTGTGGAGCTGAGTTCCGACACCTGCACTGAAAGCAAGAAA-----	1540	
OY	1326	AATGAGGCGGTGCCAATTTGATGCTCTCTTGTATGTGATGAGAGAGCTGCACTATC	1385	
Db	1541	AAGCTTGGAAACAGAACTAATGAGGGGCTCTCATTTGACAGAAACTTCACTCTCT	1600	
OY	1386	ACCTTGAGACTGAGTGTACCAAGGCTCAAGATGAGACCTAGAGACCCACTCTTG	1445	
Db	1601	AGCTTGAACCCAGTTGTGCGACCGAGGCTTGGTGTATGACTGGAGACCACTCTCT	1660	
OY	1446	CCAGTTGGTGTATCTCCAACTGTGCATGTGCAATGCGCAATGCTTGGCATCACTCTG	1505	
Db	1661	CTGTGCTGTGATCTCCAGCTGACCGACGCTCCCGATGGCTGGCGCTTATCTGTGG	1720	
OY	1506	TATACATGCTGACCAATTAACCCAGAACGTGAAGCTTCTTCACTAAGCCGCCAATTG	1565	
Db	1721	TACAAACATGCTGTGAGAGAGCCAGGAATCTCTCTTCTTGAACCCCGCTGGCGC	1780	
OY	1566	ACCTGGAGCAAGTGGCCGAGAGTGTCAAGCTGGAGTTCTGTCCACCAAGCAGAGG	1625	
Db	1781	TGGTGTCCCGCTCTAGAGAGTGTGAGTGAAGTTTATGATGATCAAGCAAGAGAGT	1840	
OY	1626	CTGAGCATGAGCAGCTGACAAAGCTGTGAGAAAGCTCTAGAGGCTGTGTGTAATAC	1685	
Db	1841	CTTGAACGAGACCAAGTGTGATGCTGGAGAGAAAGTGTGGGCTTAATGC---TGCG	1897	
OY	1686	TCAGGTGTGATGATCAATGAGCTAAATTTGCAAAAGAAACATGAGCTGGCAGAGGCT	1745	
Db	1898	CCTGATGCTTATTCATTCAGACAGGTTTGTAGGAAATATTAATGATTAATTAATTT	1957	
OY	1746	TCCTTGTGGTGTGCTTGAACATATCATGCACTTGTGAAAAATATATCTGTGCCCT	1805	
Db	1958	TCCTTGTGGCTTGTGATGACACCATCTCTAGAGCTTATTAAGAACACCTGTGTGCTG	2017	
OY	1806	TGGAATGAAGGTATCATGTGCTTTCATTCAGCAAGAGCGGAGCGGCAATCTCAAG	1865	
Db	2018	TGGAATGATGGGTGATTAATGAGCTTTCATCAGCAAGAGCGGAGCGGCTGTCTAAG	2077	
OY	1866	ACAAAGCCCGGCGACCTTCTTACTGCGCTTTCAGCGAGAGCAGCAAAAGAGAGGCTC	1925	
Db	2078	GACCAAGAGCGGAGAGCTTCTGCTTGAATTCAGTGAAGAGCTCCCGGAAAGGCGCATC	2137	
OY	1926	ACTTTCACCTTGGGTGGAAGAGAC---ATCAGTGGCAAGACCCACATCCAGTCTAG	1982	
Db	2138	ACATTCACATGGGTGGAAGGTCCTCAAGAGGAGGTGAACCTGACTTCATGCGCTGAG	2197	
OY	1983	CCATACCAACACAGCAGCTGCAACAATGTGATTTGCTGAATCATCATGAGCTATAG	2042	
Db	2198	CCCTACAGAAAAAAGAACTTTCAGCTGTACTTTCCAGATATTAATTCGCAACTACAAA	2257	
OY	2043	ATCATGATGCGACCAACATCTGTGTCTCCACTTGTCTACCTTACCCGACATCTCC	2102	
Db	2258	GTCATGGCTGCCGAAACATTCACAGAAATCCCTGAGATFATGTATACCAATTTGAC	2317	
OY	2103	AAGGAGAGGCAATTTGGAAGTACTGT	2129	
Db	2318	AAAGACCAAGCCCTTGGGAAGTATTAAT	2344	

RESULT 12
B1108275 719 bp mRNA EST 26-JUN-2001
LOCUS

DEFINITION 602902137f1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5032016 5', mRNA sequence.
ACCESSION B1108275
VERSION B1108275.1 GI:14559181
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 719)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1AM1089 row: 1 column: 09
High quality sequence stop: 719.
Location/Qualifiers
1..719
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5032016"
/clone_lib="NIH_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert 2 kb. Library constructed by Life Technologies, catalog #12017-018. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999). Note: this is a NCI_CGAP Library."
BASE COUNT 184 a 198 c 187 g 150 t
ORIGIN
Query Match 24.2%; Score 694.4; DB 11; Length 719;
Best Local Similarity 99.9%; Pred. No. 6,2e-149;
Matches 693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1679 GAAGTACTCAGGGGTGCAGATCATGCGCTTAATTCGCAAGAAACATGCGCTCGCA 1738
DB 24 GAAGTACTCAGGGGTGCAGATCATGCGCTTAATTCGCAAGAAACATGCGCTCGCA 83
OY 1739 GGGCTTCTCTTCTGGGCTGCTAGACAAATATCATGACCTTGTGAAAAGTATATCTT 1798
DB 84 GGGCTTCTCTTCTGGGCTGCTAGACAAATATCATGACCTTGTGAAAAGTATATCTT 143
OY 1799 GGGCTTCTCTTCTGGGCTGCTAGACAAATATCATGACCTTGTGAAAAGTATATCTT 1858
DB 144 GGGCTTCTCTTCTGGGCTGCTAGACAAATATCATGACCTTGTGAAAAGTATATCTT 203
OY 1859 CCTAAGCACAAAGCCCGGGACCTTCTACTGCGCTTCAGCGAGAGCAGCAAAAGG 1918
DB 204 CCTAAGCACAAAGCCCGGGACCTTCTACTGCGCTTCAGCGAGAGCAGCAAAAGG 263
OY 1919 AGGGGTCACTTTCAGTTGGGTGGAAGAGACATCACTGCGAAGACCCAGACTCTGT 1978
DB 264 AGGGGTCACTTTCAGTTGGGTGGAAGAGACATCACTGCGAAGACCCAGACTCTGT 323
OY 1979 AGAGCATACACCAAGCAGACAGCTGAACAACATGTCATTTGTAATCATATGCGCTA 2038
DB 324 AGAGCATACACCAAGCAGACAGCTGAACAACATGTCATTTGTAATCATATGCGCTA 383
OY 2039 TAAGTCAATGATGGACCAACATCCTGTGCTCAGTTGCTACCTTACCCGACAT 2098
|||||

DB 384 TAAGTCAATGATGGACCAACATCCTGTGCTCAGTTGCTACCTTACCCGACAT 443
OY 2099 TCCCAAGAGAGAGCATTTGGAAAGTACTGTAGGCCGAGAGCCAGACCCCGAAGC 2158
DB 444 TCCCAAGAGAGAGCATTTGGAAAGTACTGTAGGCCGAGAGCCAGACCCCGAAGC 503
OY 2159 CGACCCAGGTAGTGTGCCCCGCTAGCTGAAGACCAAGTTCATGTGTGACACCAAGC 2218
DB 504 CGACCCAGGTAGTGTGCCCCGCTAGCTGAAGACCAAGTTCATGTGTGACACCAAGC 563
OY 2219 CTCGACCAATACATTTAGCTAGCTGCGGATGTCCCCGACACTTATGATTCATGTGAGTT 2278
DB 564 CTCGACCAATACATTTAGCTAGCTGCGGATGTCCCCGACACTTATGATTCATGTGAGTT 623
OY 2279 TCGAATTAACGCTGAAGTGTCTGAGCCCTCAGCAGAGGAGCAGTTGAGTCGCTACGTT 2338
DB 624 TCGAATTAACGCTGAAGTGTCTGAGCCCTCAGCAGAGGAGCAGTTGAGTCGCTACGTT 683
OY 2339 TGACATGATCTGACCTCGAGTGTCTTACCTCCC 2374
DB 684 TGACATGATCTGACCTCGAGTGTCTTACCTCCC 719
RESULT 13
BG973911 714 bp mRNA EST 12-JUN-2001
LOCUS 602843546F1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4979458 5', mRNA sequence.
DEFINITION BG973911
ACCESSION BG973911 GI:14361548
VERSION BG973911.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus m.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 714)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1AM1097 row: k column: 11
High quality sequence stop: 714.
Location/Qualifiers
1..714
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4979458"
/clone_lib="NIH_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert 2.5 kb. Library constructed by Life Technologies, catalog # 12018-016. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996). Note: this is a NCI_CGAP Library."
BASE COUNT 184 a 187 c 195 g 148 t
ORIGIN
Query Match 24.0%; Score 688.4; DB 11; Length 714;
Best Local Similarity 99.6%; Pred. No. 1.5e-147;
|||||

Matches 711: Conservative 0: Mismatches 1: Indels 2: Gaps 2:

```

Oy 1284 GCAGAGTTAAGCACCCTGACCCCTTAGGAGCAGAGATGTGGATGAGAGCCGCTGCCAAT 1343
      |||||
      1 GCAGAGTTAAGCACCCTGACCCCTTAGGAGCAGAGATGTGGATGAGAGCCGCTGCCAAT 60
Oy 1344 TGTGATGCTCTCTTGTGCTGACTGAGAGCTGACCTGATCATCTTCGAGACTGAGGTG 1403
      |||||
      61 TGTGATGCTCTCTTGTGCTGACTGAGAGCTGACCTGATCATCTTCGAGACTGAGGTG 120
Oy 1404 TACCACCAAGGCTCCTCAAGATTGACCTAGAGACCCTGCTGGCAGTTGGTGGATCTGC 1463
      |||||
      121 TACCACCAAGGCTCCTCAAGATTGACCTAGAGACCCTGCTGGCAGTTGGTGGATCTGC 180
Oy 1464 AACATCTGTGATGCTCAATGCTTGGGATCAATCCTGTGATTAACATGCTGACCAAT 1523
      |||||
      181 AACATCTGTGATGCTCAATGCTTGGGATCAATCCTGTGATTAACATGCTGACCAAT 240
Oy 1524 AACCCCAAGACGTGAACCTCTTCACTAAGCCGCAATTGGAACTGGGAGCAAGTGGCC 1583
      |||||
      241 AACCCCAAGACGTGAACCTCTTCACTAAGCCGCAATTGGAACTGGGAGCAAGTGGCC 300
Oy 1584 GAGGTGCTGAGCTGCGAGTTCTCG-TCACACCAAGCGAGGCTGAGCATCGAG-AGCT 1642
      |||||
      301 GAGGTGCTGAGCTGCGAGTTCTCG-TCACACCAAGCGAGGCTGAGCATCGAGCATCGAGAGCT 360
Oy 1643 GACACGCTGCTGAGAGAGCTCTAGGCTGCTGCTGAGACTACTCAGAGGTGTGATGATCAG 1702
      |||||
      361 GACACGCTGCTGAGAGAGCTCTAGGCTGCTGCTGAGACTACTCAGAGGTGTGATGATCAG 420
Oy 1703 ATGGGCTAAATTTGCAAAAGAAACATGCTGCGAAGGCTTCTCTTCTGGGCTTGAGCT 1762
      |||||
      421 ATGGGCTAAATTTGCAAAAGAAACATGCTGCGAAGGCTTCTCTTCTGGGCTTGAGCT 480
Oy 1763 AGCAATATCATGAGACCTTGTGAAGATATCTTGGCCCTTTGGAGTGAAGGATACAT 1822
      |||||
      481 AGCAATATCATGAGACCTTGTGAAGATATCTTGGCCCTTTGGAGTGAAGGATACAT 540
Oy 1823 CATGGGTTTCATGAGAGAGGAGGAGGCGGCGATCTGTAAGCAAAAGCCCGGCGAC 1882
      |||||
      541 CATGGGTTTCATGAGAGAGGAGGAGGCGGCGATCTGTAAGCAAAAGCCCGGCGAC 600
Oy 1883 CTTCCTACTGCTGCTTACGAGAGAGCAAGCAAAAGAGGAGGCTACTTTCAGTTGGGTGA 1942
      |||||
      601 CTTCCTACTGCTGCTTACGAGAGAGCAAGCAAAAGAGGAGGCTACTTTCAGTTGGGTGA 660
Oy 1943 A-AAGGACATCAGTGGCAAGACCCAGATCCAGTCTGTAGAGCCATACACCAAGC 1995
      |||||
      661 A-AAGGACATCAGTGGCAAGACCCAGATCCAGTCTGTAGAGCCATACACCAAGC 714

```

RESULT 14

B1155944

LOCUS 736 bp mRNA EST 05-JUL-2001

DERIVATION 602904278F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5033955 5',

ACCESSION B1155944

VERSION B1155944.1 GI:14615957

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 736)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Lotmar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1094 row: k column: 20
 High quality sequence stop: 732.
 Location/Qualifiers

FEATURES

source

1..736

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_image="IMAGE:503395"

/clone_lib="NIH_CGAP_Mam3"

/tissue_type="tumor, gross tissue"

/lab_host="DH10B"

/note="Organ: mammary. Vector: pCMV-Sport6. Site: 1: NotI; Site: 2: SalI. Cloned unidirectionally. Primer: Oligo dT. Average insert 2 kb. Library constructed by Life Technologies, catalog #12017-018. Investigators providing samples: Lotmar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999). Note: this is a NCI_CGAP Library."

BASE COUNT 199 a 178 c 204 g 155 t
 ORIGIN

Query Match 23.7%; Score 679; DB 11; Length 736;

Best Local Similarity 98.8%; Pred. No. 2.1e-145;

Matches 726: Conservative 0: Mismatches 5: Indels 4: Gaps 4:

```

Oy 735 GGGCTCTTGTCACCAATGAGTACGTCGACAGACACTGACTGATGAAGAGCTGGCTGAC 794
      |||||
      1 GGGCTCTTGTCACCAATGAGTACGTCGACAGACACTGACTGATGAAGAGCTGGCTGAC 60
Oy 795 TGGAGAGGGGGGCGAGATCGGTCGATCGAGAGGCTCCCAACATCTGCTGAGACCT 854
      |||||
      61 TGGAGAGGGGGGCGAGATCGGTCGATCGAGAGGCTCCCAACATCTGCTGAGACCT 120
Oy 855 CTGAA-?-ACTGATACTTCATTAGCAGATCTCAACTTCAGACCCGCCAACAAATTAA 913
      |||||
      121 CTGAAACACTTGATACCTTCACTAGCAGATCTCAACTTCAGACCCGCCAACAAATTAA 180
Oy 914 GAAACGTGAGAGCTGAGAGAAAGTCTCTACAAAGGGGAGCCCTTATGTCAGACACCG 973
      |||||
      181 GAAACGTGAGAGCTGAGAGAAAGTCTCTACAAAGGGGAGCCCTTATGTCAGACACCG 240
Oy 974 GCCCATGCTGAGAGAGAGATCTGTGAGCTGTCAGAACTTAATGAAGAGTCCCTTCGT 1033
      |||||
      241 GCCCATGCTGAGAGAGAGATCTGTGAGCTGTCAGAACTTAATGAAGAGTCCCTTCGT 300
Oy 1034 GGTGAGAGGGGAGCTCCCTGTCATGCCCATGACCCGAGCCGCCCTTACTCATCAAGACTGG 1093
      |||||
      301 GGTGAGAGGGGAGCTCCCTGTCATGCCCATGACCCGAGCCGCCCTTACTCATCAAGACTGG 360
Oy 1094 TGTTCAGTTTACACGAAACTGAGTGTGCTGTCGAATTTCTGAGTTGAATTATGAGCT 1153
      |||||
      361 TGTTCAGTTTACACGAAACTGAGTGTGCTGTCGAATTTCTGAGTTGAATTATGAGCT 420
Oy 1154 TAAATTTAAAGTGTGATTTGATAAAGACTGTGGGATGTTGCTGCTCAGAGGCTTCG 1213
      |||||
      421 TAAATTTAAAGTGTGATTTGATAAAGACTGTGGGATGTTGCTGCTCAGAGGCTTCG 480
Oy 1214 G-AAATTTAATCATTTGCGGACGAGCAAAAGATGATGAGCATGAGAGATCTTAACAG 1272
      |||||
      481 GAAATTTAATCATTTGCGGACGAGCAAAAGATGATGAGCATGAGAGATCTTAACAG 540
Oy 1273 GCAAGCTGTCTGC-AGAGTTCAAGCACCCTGAGGAGCAGAGATGTGGAAATGGA 1331
      |||||
      541 GCAAGCTGTCTGC-AGAGTTCAAGCACCCTGAGGAGCAGAGATGTGGAAATGGA 600
Oy 1332 GGGCGTCCCAATTTGTATGGCTCTTGTATGCTGATGAGAGAGCTGACCTGATACCTTC 1391
      |||||
      601 GGGCGTCCCAATTTGTATGGCTCTTGTATGCTGATGAGAGAGCTGACCTGATACCTTC 660
Oy 1392 GAGACTGAGTGTACCAAGGCTCAAGATTGAGCT-AGAGACCACTCTTGGCCAGT 1450

```

```
Db 661 GAGACTGAGGTGTACACCAAGGCGCTCAAGATGACCTAAGAGAACCACTCTCCAGT 720
OY 1451 TGTGCTGATCTCCAA 1465
Db 721 ACTGCTGATCTCCAA 735

RESULT 15
B1113023 717 bp mRNA EST 26-JUN-2001
LOCUS 602898184F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5028140 5',
DEFINITION mRNA sequence.
ACCESSION B1113023 GI:14563924
VERSION B1113023.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 717)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9qbs@email.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11079 row: g column: 21
High quality sequence stop: 678.
Location/Qualifiers
1..717
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
/clone="IMAGE:5028140"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT 188 a 193 c 192 g 144 t
ORIGIN

Query Match 23.6%; Score 676.6; DB 11; Length 717;
Best Local Similarity 99.2%; Pred. No. 7.6e-145;
Matches 701; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

OY 1536 GTGACCTCTTCACTAAGCGCCCAATTGGAACCTGGACCAAGTGGCGAGGTGCTCAGC 1595
|||||
Db 1 GTGAACCTTCTTCACTAAGCGCCCAATTGGAACCTGGACCAAGTGGCGAGGTGCTCAGC 60
|||||

OY 1596 TGGCAGTCTCTGTCGACCAACGAGGCGTGAACATGAGCAGCTGACGCTGAGT 1655
|||||
Db 61 TGGCAGTCTCTGTCGACCAACGAGGCGTGAACATGAGCAGCTGAGT 120
|||||

OY 1656 GAGAGCTCTGAGGCGCTGGTGAACACTACAGGCTGTCACATGACATGGGCTAAATTC 1715
|||||
Db 121 GAGAGCTCTGAGGCGCTGGTGAACACTACAGGCTGTCACATGACATGGGCTAAATTC 180
|||||

OY 1716 TGCAAGAAAGAAACATGCTGCGAAGGCTTCTCTCTGAGTCTGCTAGACATATCATC 1775
|||||
Db 181 TGCAAGAAAGAAACATGCTGCGAAGGCTTCTCTCTGAGTCTGCTAGACATATCATC 240
|||||
```

```
OY 1776 GACCTTGTGAAAAGTATATCTTGCCCTTTGGATGAAGGCTACATCATGGTTTCATC 1835
|||||
Db 241 GACCTTGTGAAAAGTATATCTTGCCCTTTGGATGAAGGCTACATCATGGTTTCATC 300
|||||

OY 1836 AGCAAGGAGCGGAGCGGCGCATCTTAAGCACAAAGCCCCGGGACCTCTCTACTGCGC 1895
|||||
Db 301 AGCAAGGAGCGGAGCGGCGCATCTTAAGCACAAAGCCCCGGGACCTCTCTACTGCGC 360
|||||

OY 1896 TTTACGAGAGCAGCAAGAAAGAGGGTCACTTTCACTTGGGTGAAAAGACATCAGT 1955
Db 361 TTCAGCGAATACGAGAAAGAAAGAGGGGTCACTTTCACTTGGGTGAAAAGACATCAGT 420

OY 1956 GGCAGACCCAGATCCAGTCTGTAGAGCATACAGCAAGCAGCAGCTGTAACAACATGCA 2015
|||||
Db 421 GGCAGACCCAGATCCAGTCTGTAGAGCATACAGCAAGCAGCAGCTGTAACAACATGCA 480

OY 2016 TTTGCTGAATCATCATGGGTATAGATCAT -GGATGCGACCAACATCTGTGTCCTC 2074
|||||
Db 481 TTTGCTGAATCATCATGGGTATAGATCATGGATGGACCAACATCTGTGTCCTC 540

OY 2075 ACTTGCTACCTCTACCCCGACATTCCCAAGSAGAGGATTTGGAAGTACTGTAGCC 2134
|||||
Db 541 ACTTGCTACCTCTACCCCGACATTCCCAAGSAGAGGATTTGGAAGTACTGTAGCC 600

OY 2135 CGAGAGCTGAGAGCACCCCGAAGCCGACCCAGTAGTGCTGCCCGTACCTGAAGACCAA 2194
|||||
Db 601 CGAGAGCTGAGAGCACCCCGAAGCCGACCCAGTAGTGCTGCCCGTACCTGAAGACCAA 660

OY 2195 GTTCATCTGTGTGACACCAACGACCTGCAACATACATTGACCTGC 2241
|||||
Db 661 GTTCATCTGTGTGACACCAACGACCTGCAACATACATTGACCTGC 706
|||||
```

Search completed: March 19, 2002, 14:30:29
Job time: 1772.0000

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:17:33 ; Search time 33.38 Seconds
(without alignments)
1757.172 Million cell updates/sec

Title: US-08-212-185-12

Perfect score: 40729
Sequence: 1 MAQMNQLOQLNTRYLKQLHQ.....QFESLTFMDLTSKATSPM 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4006	99.4	770	2	I49508	ISGF3 p91-related
2	3977	98.7	770	2	A54444	DNA-binding protei
3	1971.5	48.9	739	2	A46159	interferon-depende
4	1782	44.2	748	2	A56047	gamma-interferon a
5	1293	32.1	851	2	A46160	interferon alpha-1
6	869.5	21.6	794	2	G02317	transcription acti
7	868.5	21.6	786	2	I49274	mammary gland fact
8	858	21.3	793	2	S54772	mammary gland fact
9	791.5	19.6	794	2	S55527	mammary gland fact
10	582	14.4	837	2	I57557	DNA-binding protei
11	555.5	13.8	848	2	A54740	interleukin-4-indu
12	164.5	4.1	1356	2	S32763	kinectin 1 - huma
13	159.5	4.0	2101	2	A42184	nuclear mitotic ap
14	159	3.9	1085	2	JC2227	probable helicase
15	156.5	3.9	1300	2	I53799	CG1 protein - huma
16	152.5	3.8	64	2	S63681	signal transducer
17	152	3.8	978	2	A70387	conserved hypochet
18	148	3.7	65	2	S63679	signal transducer
19	148	3.7	2663	1	S28261	centromere protein
20	147	3.6	1410	1	A57013	early endosome ant
21	145	3.6	1427	2	S22695	reslin - human
22	145	3.6	1790	2	S67593	transport protein
23	144.5	3.6	1959	1	A33977	myosin heavy chain
24	144	3.6	876	2	A23767	myosin heavy chain
25	144	3.6	1006	2	S74992	hypothetical prote
26	144	3.6	1780	2	T17272	hypothetical prote
27	141.5	3.5	924	2	C86725	chromosome segrega
28	141	3.5	734	2	T27055	hypothetical prote
29	141	3.5	741	2	S39082	myosin heavy chain

30	141	3.5	1922	2	T00637	hypothetical prote
31	140.5	3.5	1312	1	BMBYDL	RAD50 protein - ye
32	140.5	3.5	1526	2	T41522	myosin II - fistro
33	140.5	3.5	1690	2	T13030	microtubule bindin
34	140.5	3.5	1738	2	T14867	interactin - slime
35	140	3.5	936	2	S39083	myosin heavy chain
36	140	3.5	955	2	S24348	myosin heavy chain
37	140	3.5	1166	2	T27075	hypothetical prote
38	140	3.5	1940	2	A29320	myosin heavy chain
39	140	3.5	3685	1	A27605	dystrophin, muscl
40	139	3.4	1188	2	G83960	chromosome segrega
41	138.5	3.4	2168	2	T30171	ninein - mouse
42	138	3.4	1642	2	T08880	NMDA receptor-biu
43	138	3.4	2385	2	A32491	myosin heavy chain
44	138	3.4	2411	2	B32491	myosin heavy chain
45	137.5	3.4	1313	2	F96673	hypothetical prote

ALIGNMENTS

RESULT	1	
I49508		
ISGF3 p91-related transcription factor - mouse		
C:Species: Mus musculus (house mouse)		
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000		
C:Accession: I49508; I49009		
R:AKira, S., Nishio, Y., Inoue, M., Wang, X.J., Wei, S., Matsusaka, T., Yoshida, K.,		
Cell 77, 63-71, 1994		
A:Title: Molecular Cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related		
A:Reference number: A54444; MIMD:94208062		
A:Accession: I49508		
A>Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: mRNA		
A:Residues: 1-770 <RES>		
A:Cross-references: GB:I29278; NID:9476715; PIDN:AAA37254.1; PID:9476716		
R:Raz, R., Durbin, J.E., Levy, D.E.		
J. Biol. Chem. 269, 24391-24395, 1994		
A:Title: Acute phase response factor and additional members of the interferon-stimula		
A:Reference number: I49009; MIMD:95014185		
A:Accession: I49009		
A>Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: mRNA		
A:Residues: 1-393; M:395-700; 702-770 <RE2>		
A:Cross-references: EMBL:U08378; NID:9473889; PIDN:AAA56668.1; PID:9473890		
C:Genetics:		
A:Gene: APRF		
C:Superfamily: human signal transducer and transcription activator STAT5A		
Query Match	99.4%;	Score 4006; DB 2; Length 770;
Best Local Sim. only	99.4%;	Pred. No. 1.2e-239;
Matches 765	1; Mismatches 4;	Indels 0; Gaps 0;
Db	1	MAQMNQLOQLNTRYLKQLHQ.....QFESLTFMDLTSKATSPM 770
Qy	1	MAQMNQLOQLNTRYLKQLHQ.....QFESLTFMDLTSKATSPM 770
Db	1	MAQMNQLOQLNTRYLKQLHQ.....QFESLTFMDLTSKATSPM 770
Qy	61	LGELIDQYSRFLQESNVLYOHNLRIKQFLQSRLEKPMELIARIVACILWEESRLQTA 120
Db	61	LGELIDQYSRFLQESNVLYOHNLRIKQFLQSRLEKPMELIARIVACILWEESRLQTA 120
Qy	121	TAAQOCGQANHPPTAAVTEKQOMLEQHLQDVRRKQVQDLQKMKVENVLQDDPFNNKTLK 180
Db	121	TAAQOCGQANHPPTAAVTEKQOMLEQHLQDVRRKQVQDLQKMKVENVLQDDPFNNKTLK 180
Qy	181	SQGMQDINGNOSVTRKQKQOLEQMLTALDOMBRSTVSELAGLSMKEYVQKTLDEEL 240
Db	181	SQGMQDINGNOSVTRKQKQOLEQMLTALDOMBRSTVSELAGLSMKEYVQKTLDEEL 240
Qy	241	ADMKRREIACIGGPNICIDRLLENMTTSLAESQLQTRQOIKKLEELQKQVSKGPPIVQ 300
Db	241	ADMKRREIACIGGPNICIDRLLENMTTSLAESQLQTRQOIKKLEELQKQVSKGPPIVQ 300

QY 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPRPLVKTGVQFTTKVRLLYKPELANT 360
|||||
Db 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPRPLVKTGVQFTTKVRLLYKPELANT 360
QY 361 QLTAKVCIDKDSGVAAALRGRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREORCGN 420
|||||
Db 361 QLTAKVCIDKDSGVAAALRGRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREORCGN 420
QY 421 GGRANCASLIVTEELHLITFETEVYHQGLKIDLETHSLPVVVISNICOMPMAASILMY 480
|||||
Db 421 GGRANCASLIVTEELHLITFETEVYHQGLKIDLETHSLPVVVISNICOMPMAASILMY 480
QY 481 NMLTNKRNKNNVFTKPPIGTMDOVAEVLWQFSSSTTKRGSLIEOLTLAKELGPGVNS 540
|||||
Db 481 NMLTNKRNKNNVFTKPPIGTMDOVAEVLWQFSSSTTKRGSLIEOLTLAKELGPGVNS 540
QY 541 GCOTTMKFCCKENNAKGFSFWMLDNIIDLKYYILALMNEGYIMGFISKEERAILST 600
|||||
Db 541 GCOTTMKFCCKENNAKGFSFWMLDNIIDLKYYILALMNEGYIMGFISKEERAILST 600
QY 601 KPCTFLFRSESSKEGCVFTFWWEKDISGKTQIOSVEPYTKOOLNNMFAEIIIMGYKIM 660
|||||
Db 601 KPCTFLFRSESSKEGCVFTFWWEKDISGKTQIOSVEPYTKOOLNNMFAEIIIMGYKIM 660
QY 661 DATNILVSPLYLYLPDIPKEAFKCYCRPESQEHPEADPGSAAPYLKTKFICVTPTTCSN 720
|||||
Db 661 DATNILVSPLYLYLPDIPKEAFKCYCRPESQEHPEADPGSAAPYLKTKFICVTPTTCSN 720
QY 721 TIDLPMSPRTDLSLMQFGNNGEGAPESAGQFESLTFDMDLTSCATSPM 770
|||||
Db 721 TIDLPMSPRTDLSLMQFGNNGEGAPESAGQFESLTFDMDLTSCATSPM 770
RESULT 2
DNA-binding protein APRF - human
C:Species: Homo sapiens (man)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 28-Jul-2000
C:Accession: A54444
R:Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Sue
Cell 77, 63-71, 1994
A:Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra
A:Reference number: A54444; MUID:94208062
A:Accession: A54444
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-770 <RES>
A:Cross-references: GB:I29277; NID:q475788; PID:q475789
C:Genetics:
A:Gene: GDB:STAT3; APRF
A:Cross-references: GDB:358950
A:Map position: 17q21-17q21
C:Superfamily: human signal transducer and transcription activator STAT5A
C:Keywords: DNA binding; transcription factor

Query Match 98.7%; Score 3977; DB 2; Length 770;
Best Local Similarity 98.6%; Pred. No. 7,le-238;
Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAOMNLOOUDTRYLKOLHOLYSDFPMELOFLAPWIESODMAVAASKESHATLVFHNL 60
|||||
Db 1 MAOMNLOOUDTRYLKOLHOLYSDFPMELOFLAPWIESODMAVAASKESHATLVFHNL 60
QY 61 LGEIDQYSRFLQESNVLYOHNLRRIKQFLOSRYLEKPMETIARIVARCLMEESRLIQTAA 120
|||||
Db 61 LGEIDQYSRFLQESNVLYOHNLRRIKQFLOSRYLEKPMETIARIVARCLMEESRLIQTAA 120
QY 121 TAAOOGGANHTAATVTEKQOMLEOHLODVARKRVODLEOKMKVVENLODPDFNYKTLK 180
|||||
Db 121 TAAOOGGANHTAATVTEKQOMLEOHLODVARKRVODLEOKMKVVENLODPDFNYKTLK 180

QY 181 SOGDMODLNNOSVTRQKMOLEOMHTALDOMRSTIVSELAGLLSMEVYOKLTDEEL 240
|||||
Db 181 SOGDMODLNNOSVTRQKMOLEOMHTALDOMRSTIVSELAGLLSMEVYOKLTDEEL 240
QY 241 ADMKRPPEIACIGPNNICIDRLLENMTTSLAESOLQTRQOIKKLEELQOKVSKGDPYVQ 300
|||||
Db 241 ADMKRPPEIACIGPNNICIDRLLENMTTSLAESOLQTRQOIKKLEELQOKVSKGDPYVQ 300
QY 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPRPLVKTGVQFTTKVRLLYKPELANT 360
|||||
Db 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPRPLVKTGVQFTTKVRLLYKPELANT 360
QY 361 QLTAKVCIDKDSGVAAALRGRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREORCGN 420
|||||
Db 361 QLTAKVCIDKDSGVAAALRGRKFNILGTNTKVMNEESNNGSLSAEFKHLTLREORCGN 420
QY 421 GGRANCASLIVTEELHLITFETEVYHQGLKIDLETHSLPVVVISNICOMPMAASILMY 480
|||||
Db 421 GGRANCASLIVTEELHLITFETEVYHQGLKIDLETHSLPVVVISNICOMPMAASILMY 480
QY 481 NMLTNKRNKNNVFTKPPIGTMDOVAEVLWQFSSSTTKRGSLIEOLTLAKELGPGVNS 540
|||||
Db 481 NMLTNKRNKNNVFTKPPIGTMDOVAEVLWQFSSSTTKRGSLIEOLTLAKELGPGVNS 540
QY 541 GCOTTMKFCCKENNAKGFSFWMLDNIIDLKYYILALMNEGYIMGFISKEERAILST 600
|||||
Db 541 GCOTTMKFCCKENNAKGFSFWMLDNIIDLKYYILALMNEGYIMGFISKEERAILST 600
QY 601 KPCTFLFRSESSKEGCVFTFWWEKDISGKTQIOSVEPYTKOOLNNMFAEIIIMGYKIM 660
|||||
Db 601 KPCTFLFRSESSKEGCVFTFWWEKDISGKTQIOSVEPYTKOOLNNMFAEIIIMGYKIM 660
QY 661 DATNILVSPLYLYLPDIPKEAFKCYCRPESQEHPEADPGSAAPYLKTKFICVTPTTCSN 720
|||||
Db 661 DATNILVSPLYLYLPDIPKEAFKCYCRPESQEHPEADPGSAAPYLKTKFICVTPTTCSN 720
QY 721 TIDLPMSPRTDLSLMQFGNNGEGAPESAGQFESLTFDMDLTSCATSPM 770
|||||
Db 721 TIDLPMSPRTDLSLMQFGNNGEGAPESAGQFESLTFDMDLTSCATSPM 770
RESULT 3
A46159
Interferon-dependent positive-acting transcription factor ISGF-3 91k chain - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Jul-2000
C:Accession: A46159
R:Schindler, C.; Fu, X.Y.; Impirota, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
A:Title: Proteins of transcription factor ISGF-3: one gene encodes the 91- and 84-kDa
A:Reference number: A46159; MUID:92366557
A:Accession: A46159
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-739 <SCH>
A:Experimental source: Hela cells
A:Note: Sequence extracted from NCBI backbone (NCBIP:110818)
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 48.9%; Score 1971.5; DB 2; Length 739;
Best Local Similarity 51.7%; Pred. No. 4,7e-114;
Matches 383; Conservative 144; Mismatches 189; Indels 25; Gaps 10;

QY 1 MAOMNLOOUDTRYLKOLHOLYSDFPMELOFLAPWIESODMAVAASKESHATLVFHNL 60
|||||
Db 1 MSOWEELQOOLDSKLEOVHOLYSDSPMEIROYLQWOLEKODMAHANDVSFAIRFHD 60
QY 61 LGEIDQYSRFLQESNVLYOHNLRRIKQFLOSRYLEKPMETIARIVARCLMEESRLIQTAA 120
|||||
Db 61 LSQIDQYSRFLQESNVLYOHNLRRIKQFLOSRYLEKPMETIARIVARCLMEESRLIQTAA 120
QY 121 TAAOOGGANHTAATVTEKQOMLEOHLODVARKRVODLEOKMKVVENLODPDFNYKTLK 180
|||||


```

Db      121 RFND--AOSGNIGSTVMDLKKOELDSKVKNYKDKMCEHEIKSLIEDLQDEYDKCKTL- 177
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy      181 SOGDMDLNGNNOVSTROKMOOLEOMLTALDOMRSIVSELAGLSAMEYVOKTLTDEEL 240
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      178 -QNRHEHNGVAKSQOQKDOULLKMKMYLMDLNKREYVHKIIELLNVTELRQNALINDEL 236
Oy      241 ADMKRRPEIACIGPPNICLOLRLENMTISLAEOLQTRQOIKKLEELQOKYSKCDPIVQ 300
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      237 VEMKRRQOSACIGPPNACLDOLO-----QVRQOLKKLEELQOKYTEHDPIYTK 285
Oy      301 HRPMLLEERIVELFRNLMSAFVEROPCPMPHDPRLVTKIGVOPFTTKVRLVLPKPELNY 360
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      286 NKQVLMDRFTSLFOOLIOSSFVEROPCPMPHDPRLVTKIGVOPFTTKVRLVLPKPELNY 345
Oy      361 OLKIKVCIKDSGDVAALRGSKRNILGNTKVMNMEESNNGSLAEFKHILTLREORCN 420
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      346 NLKVKVLEFDKDVNERNTYVGRFKNILGHTKVMNMEESTNGSLAEFRHLQLEOK--N 403
Oy      421 GGRANDASLIYTEELHLITFEFEYVHOGKIDLETHSLPVVAVINICOMPAMASILMY 480
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      404 ACTRTNCEPLIYTEELHLSFETOLCOPGLVIDLETSLPVVAVINISOLPSGMAASILMY 463
Oy      481 NMLTNPKNVNFFTKPPICGTMDQVAEVLSMQFSSTTKRGLSIEOLTTLAEKLLGPGVNS 540
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      464 NMLVAEPNLSFELPPCARMAQLSEVLMSQFSVTKRGLNDQNLMLGEKLLGNASPD 523
Oy      541 GCOITMAKCKENMAGKGSFWVWLDNIIDLYKKYTLALMNGCYIMGFISKREBAILST 600
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      524 GL-IPMPTRCKENINDKNPFWLMIESTLEILKIKHLLPLMNDGIMGFISKREBAILSD 582
Oy      601 KPPTGFLRFSSSEKSGVTFTWVEKDLS-GKTOIOSVEPYTKOOLNNMSEAEIIMGYKI 659
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      583 QOPGFTFLRFSSSRGCAITFTWVERSONGEPDFAVAPYTKKLSAATFPDIIIRNVAV 642
Oy      660 MDATNIVSPLYLVLPDIKEPAFGK-YCRP-ESOEHPADPGSAAPYLKTKFTICVT--- 714
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      643 MAENIPENPLKLYLYPNIDKHAFCGYKSPKAEDEPMELDGPKGTGYIKTELISVSEVA 702
Oy      715 PTTCSENTID-LPMSPRTLDSL 734
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      703 PSRLQTTDNLPMSPREEPDEV 723

RESULT 4
A56047
gamma-interferon activation site-binding protein Stat4 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 28-Jul-2000
C:Accession: A56047
R:Yamamoto, K.; Queille, F.W.; Thierfelder, W.E.; Kreider, B.L.; Jenkins,
Mol. Cell. Biol. 14, 4342-4349, 1994
A:Title: Stat4, a novel gamma interferon activation site-binding protein expressed in ea
A:Reference number: A56047; MUID:94277038
A:Accession: A56047
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-748 <YAN>
A:Cross-references: GB:009351; NID:9509502; PIDN:AAA19692.1; PID:9509503
C:Superfamily: human signal transducer and transcription activator STAT5A
C:Keywords: DNA binding; phosphoprotein

```

Query Match 44.2%; Score 1782; DB 2; Length 748;
 Best Local Similarity 48.1%; Pred. No. 2,4e-102;
 Matches 356; Conservative 147; Mismatches 207; Indels 30; Gaps 11;

```

Oy      1 MAQMNLOQLDTRYLKQLQLVSDTPRMELROFLAPWISSQMAVAAKSESHATLVFNL 60
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 MSQMNVOQLDKLEFQVQDFDNDPMEIRHLAQWITQDWEVAAKSNMETATILLNL 60
Oy      61 LGEIDQVSRFLQESNVLYQHNLRKQFLOGSRYLEKPMELIRIYARCLMEESRLQTA 120
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db      61 LIQDFQILGVKSKENLLILHLKRIKRVLOQKFGHGNPMHVAIVISNCLREBRIL-AA 119
Oy      121 TAAQGGQAHNPAAVYTEKQOOMEHLQDYRRVQDLEQKMKVVELODDPDPFNKTK 180
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      120 NMPLOGLPKSLQSSVSSEQRQNRNEHKVSAIKNSVOMTEODPTKLELODEFPYRKTIO 179
Oy      181 S--OGDMDLNGNNOVSTROKMOOLEOMLTALDOMRSIVSELAGLSAMEYVOKTLTDE 238
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      180 TMDGD-----KSLVNOEVLTLQELNLSLDPKRRKDALSKMQLVNETDILMNSMLLE 233
Oy      239 ELADMKRRPEIACIGPPNICLOLRLENMTISLAEOLQTRQOIKKLEELQOKYSKCDPI 298
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      234 ELQDMKRRQOIAICIGPLHNGIDLOKCFLLAESLQLRQOLEKLOEOSTKMYTEGDP 293
Oy      299 VQHRPMLLEFRIVELFRNLMSAFVEROPCPMPHDPRLVTKIGVOPFTTKVRLVLPKPEL 358
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      294 PAQRHMLLEKRTFTLYLVLFKNSFEVEROPCPMPHDPRLVTKIGVOPFTTKVRLVLPKPEL 353
Oy      359 NYOLKIKVCIKDSGDVAALRGSKRNILGNTKVMNMEESNNGSLAEFKHILTLREORC 418
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      354 NYQVKVKSJLQK---NVSTL-SNRRFVLCGTHVAKMSSESSNGSLSEVFRHLQPKEMKC 409
Oy      419 GNGRANDASLIYTEELHLITFEFEYVHOGKIDLETHSLPVVAVINISOLPSGMAASIL 478
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      410 STGSKGN-EGCHMVTVEELHSTFETOLCPLVGLTINLETSSLPVVAVINISOLPSGMAASIL 468
Oy      479 WYNMLTNPKNVNFFTKPPICGTMDQVAEVLSMQFSSTTKRGLSIEOLTTLAEKLLGPGVN 538
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      469 WYNSTNSOHLVFPNPNPSTVGLQLFVMSQFSYVGRGLNSEQLMLAEKLTIVQS-N 527
Oy      539 YSGQITMAKCKENMAGKGSFWVWLDNIIDLYKKYTLALMNGCYIMGFISKREBAIL 598
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      528 YNDGHLTMAKCKEHLDPKGTFTFWTLEAIIIDLKIKHLLPLMIDGYIMGFISKREKHL 587
Oy      599 STKPPGFLRFSSSEKSGVTFTWVEKDLSGKTOIOSVEPYTKOOLNNMSEAEIIMGYKI 658
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      568 KDKMPGFTFLRFSSSHL-GGITFTWVDSENGEVAFHSEVYPMKGRLSALFPADILDYK 646
Oy      659 IMDATNIVSPLYLVLPDIKEPAFGK-YCRP-ESOEHPADPGSAAPYLKTKFTICVT 714
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      647 YJMAENIPENPLKLYLYPNIDKHAFCGYKSPKAEDEPMELDGPKGTGYIKTELISV 702
Oy      715 PTTCSENTID-----LPMSP 728
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      703 TIRSDSTERQSPSDDLPMSP 722

RESULT 5
A46160
Interferon alpha-induced transcription activator ISGF-3, 113k chain - human
N:Alternate names: stat2 protein
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Aug-2000
C:Accession: A46160; S71908; S53873
R:Fu, X.Y.; Schindler, C.; Impirota, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7840-7843, 1992
A:Title: The proteins of ISGF-3, the interferon alpha-induced transcriptional activat
A:Reference number: A46160; MUID:92365558
A:Accession: A46160
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA; protein
A:Residues: 1-851 <FU>
A>Note: sequence extracted from NCBI backbone (NCBIP:110820)
R:Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.
submitted to the EMBL Data Library, December 1994
A:Reference number: S71908
A:Accession: S71908
A:Molecule type: DNA
A:Residues: 1-851 <YAN>
A:Cross-references: EMBL:U18671; NID:91293919; PIDN:AAA98760.1; PID:91293920
R:Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.
Nucleic Acids Res. 23, 459-463, 1995
A:Title: The genomic structure of the STAT genes: multiple exons in coincident sites

```



```

OY 177 -KTLSQGMODLNGN-----OSVTRKMOOLEOML-----TALDOMRSISVELAGL 224
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 177 QESLRIOAFOAQLANPQERMSRETAQKQVSLFTWLQREAOITLOQYRVELAEKHOKT 236

OY 225 LLSAMEYVQKTLDEBLADKKRRPELACIGGPPNICLDRLENMTLSAESLOTROOKKL 284
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 225 LLSAMEYVQKTLDEBLADKKRRPELACIGGPPNICLDRLENMTLSAESLOTROOKKL 284

OY 237 LQLLRKQOITLIDDELQMRKROOLAGNGPREGSLDVLSWCEKLAELIWMONROOIRRA 296
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 237 LQLLRKQOITLIDDELQMRKROOLAGNGPREGSLDVLSWCEKLAELIWMONROOIRRA 296

OY 285 EELQOKVSYKGDPIVOHRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVIKTGV 344
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 285 EELQOKVSYKGDPIVOHRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVIKTGV 344

OY 297 EHLQOQOLPIRG-PVEEMLAEVNATITDIIISALVSTFIEIKP-----POLKKTQTK 347
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 297 EHLQOQOLPIRG-PVEEMLAEVNATITDIIISALVSTFIEIKP-----POLKKTQTK 347

OY 345 FTTKRVLLVFPPELNTOL---KIKVCIDKDSGVAAALRGSRKFN-----ILGTNTKVMN 396
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 345 FTTKRVLLVFPPELNTOL---KIKVCIDKDSGVAAALRGSRKFN-----ILGTNTKVMN 396

OY 348 FAATVRLVVG-GRKLNVMNPPQVATIISEOAKSLKKNENTRNECSGEL-NNCCVMEY 405
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 348 FAATVRLVVG-GRKLNVMNPPQVATIISEOAKSLKKNENTRNECSGEL-NNCCVMEY 405

OY 397 EESNNGSLSAEFKHLTREORCGNGRANCDAALVTEELHLITFETE--VYHOGKIDL 454
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 397 EESNNGSLSAEFKHLTREORCGNGRANCDAALVTEELHLITFETE--VYHOGKIDL 454

OY 406 HQA-TGTLNHFRRMSLKRKIK-----RADRRGASVTEEKFTVLFEQSFVSGNELVFOY 459
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 406 HQA-TGTLNHFRRMSLKRKIK-----RADRRGASVTEEKFTVLFEQSFVSGNELVFOY 459

OY 455 ETHSLPVPVVISNICOMPANASILMYMLTNKNVNFTRKPRIGTDQYAEVLISQFSS 514
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 455 ETHSLPVPVVISNICOMPANASILMYMLTNKNVNFTRKPRIGTDQYAEVLISQFSS 514

OY 460 KTLSLPVAVVIVHGSQDNNAATATVLMDNAFA-EPGRVP-FAVPDKVLMPOLCEALNMKEKA 517
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 460 KTLSLPVAVVIVHGSQDNNAATATVLMDNAFA-EPGRVP-FAVPDKVLMPOLCEALNMKEKA 517

OY 515 -TTKRGSLTEQTLTTLAEKLLGPGVN-----YSGCQITMAKFCENMAGKGFSPWVLNDI 568
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 515 -TTKRGSLTEQTLTTLAEKLLGPGVN-----YSGCQITMAKFCENMAGKGFSPWVLNDI 568

OY 518 EYOSNRGLTKENLFLAOKLFPNISSNLELDYNSMSVMSQFNRENLPGMNTFTWQMFQGV 577
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 518 EYOSNRGLTKENLFLAOKLFPNISSNLELDYNSMSVMSQFNRENLPGMNTFTWQMFQGV 577

OY 569 IDLVKKTALANEGYIMGFTSKERERAILSTKPRGTFLLRFSSSEKGVTFWTVVERKDI 628
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 569 IDLVKKTALANEGYIMGFTSKERERAILSTKPRGTFLLRFSSSEKGVTFWTVVERKDI 628

OY 578 MEVLKHKHRRHMNDGAILGFVNKQOAHDLINKPDCTFLFRSD-SEIGGITIAW--KFD 634
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 578 MEVLKHKHRRHMNDGAILGFVNKQOAHDLINKPDCTFLFRSD-SEIGGITIAW--KFD 634

OY 629 SGKTQIOSVEPYTKOOLNNMSFAELIMGYKIMDATNILVSPLYLYDPIDKEEAFGKYCR 688
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 629 SGKTQIOSVEPYTKOOLNNMSFAELIMGYKIMDATNILVSPLYLYDPIDKEEAFGKYCR 688

OY 635 SPDRNLMLNPKFTTRDPSISLAD-----RLGD-----LNTLIVFPDRKDEVFATYTT 684
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 635 SPDRNLMLNPKFTTRDPSISLAD-----RLGD-----LNTLIVFPDRKDEVFATYTT 684

OY 689 PE-----SQEHPE-----ADPGSAAPYLKTKFCVTPPTCSNTIDLPMSPT 730
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 689 PE-----SQEHPE-----ADPGSAAPYLKTKFCVTPPTCSNTIDLPMSPT 730

OY 685 PVLAKAVDGVKPKQKVPEFVNASTDAGASATYMDQ---APSPVACQPQ-HYNNMPPN 740
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 685 PVLAKAVDGVKPKQKVPEFVNASTDAGASATYMDQ---APSPVACQPQ-HYNNMPPN 740

OY 731 LDSLMOFGNCGAEPASAGOFESLTFPMDLTSECA 766
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 741 PDPVLD--QDGE-----FDLDESDMYA 760

RESULT 9
S55527
mammary gland factor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 28-Jul-2000
C:Accession: S55527; S44353
R:Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 14, 854-855, 1995
A:Title: Corrigenda. Mammary gland factor (MGF) is a novel member of the cytokine regula
A:Reference number: S55527; MUID:95188889
A:Accession: S55527
A:Molecule type: mRNA
A:Residues: 1-794 <WAK>
A:Cross-references: EMBL:X78428; NID:9602354; PIDN:CA55191.1; PID:9602355
A:Note: this is a revision to the sequence from reference S44353
R:Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 13, 2182-2191, 1994
A:Title: Mammary gland factor (MGF) is a novel member of the cytokine regulated transcr
A:Reference number: S44353; MUID:94244619
A:Accession: S44353
A:Molecule type: mRNA
A:Residues: 1-716; 'RHLDGGSLSR', 729, 'P', 731, 'ASL' <WAK>
A:Cross-references: EMBL:X78428
A:Note: this sequence has been revised in reference S55527
C:Superfamily: human signal transducer and transcription activator STAT5A

```

```

Query Match 19.68; Score 791.5; DB 2; Length 794;
Best Local Similarity 29.38; Pred. No. 3; 9e-41;
Matches 240; Conservative 146; Mismatches 323; Indels 109; Gaps 31;

OY 1 MAOMNLOQLDTRYLKLOHLYSDTFPMLRQFLAPPIESODMAYV-----ASKESHATLV 56
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 MACHIAQOQLQALQKMOVLYGQHPFIEVRNHLAQMTIESQPDADIDJLNPORAOYTOL 60

OY 57 FNNLGEIDQYSRPIQESNVLYOHLN-RRIKQFLQSRVLEKPEIARIVAKCLMEESRL 115
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 57 FNNLGEIDQYSRPIQESNVLYOHLN-RRIKQFLQSRVLEKPEIARIVAKCLMEESRL 115

OY 61 LEGIVQELQKAEHQVQEDSEFLIKLGHVHVHSSRTTAAFPWSMLKIRHILYNORL 120
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 LEGIVQELQKAEHQVQEDSEFLIKLGHVHVHSSRTTAAFPWSMLKIRHILYNORL 120

OY 116 LQTAATAQGGQGANHTAAVTEKQOMLEOHLDQVRRVQDLEOKKVENLQDDPDFN 175
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 116 LQTAATAQGGQGANHTAAVTEKQOMLEOHLDQVRRVQDLEOKKVENLQDDPDFN 175

OY 121 VREATNGNSSAG-----LVDAMSQKILQIQTFEELVYQDTEENELKIQOIEFIIQ 176
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 VREATNGNSSAG-----LVDAMSQKILQIQTFEELVYQDTEENELKIQOIEFIIQ 176

OY 176 Y-KTLSQGMODLNGN-----OSVTRKMOOLEOML-----TALDOMRSISVELAG 223
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 176 Y-KTLSQGMODLNGN-----OSVTRKMOOLEOML-----TALDOMRSISVELAG 223

OY 177 QESLRIOAFOAQLANPQERLSRETAQKQVSLFTWLQREAOITLOQYRVELAEKHOK 236
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 177 QESLRIOAFOAQLANPQERLSRETAQKQVSLFTWLQREAOITLOQYRVELAEKHOK 236

OY 224 LLSAMEYVQKTLDEBLADKKRRPELACIGGPPNICLDRLENMTLSAESLOTROOKKL 283
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 224 LLSAMEYVQKTLDEBLADKKRRPELACIGGPPNICLDRLENMTLSAESLOTROOKKL 283

OY 237 LQLLRKQOITLIDDELQMRKROOLAGNGPREGSLDVLSWCEKLAELIWMONROOIRR 295
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 237 LQLLRKQOITLIDDELQMRKROOLAGNGPREGSLDVLSWCEKLAELIWMONROOIRR 295

OY 284 LEELOQVSYKGDPIVOHRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVIKTGV 343
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 284 LEELOQVSYKGDPIVOHRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVIKTGV 343

OY 296 AEHLQOQOLPIRG-PVEEMLAEVNATITDIIISALVSTFIEIKP-----POLKKTQTK 346
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 296 AEHLQOQOLPIRG-PVEEMLAEVNATITDIIISALVSTFIEIKP-----POLKKTQTK 346

OY 344 QFTTKRVLLVFPPELNTOL---KIKVCIDKDSGVAAALRGSRKFN-----ILGTNTKVMN 395
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 344 QFTTKRVLLVFPPELNTOL---KIKVCIDKDSGVAAALRGSRKFN-----ILGTNTKVMN 395

OY 347 KFATVRLVVG-GRKLNVMNPPQVATIISEOAKSLKKNENTRNECSGEL-NNCCVME 404
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 347 KFATVRLVVG-GRKLNVMNPPQVATIISEOAKSLKKNENTRNECSGEL-NNCCVME 404

OY 396 MEESNNGSLSAEFKHLTREORCGNGRANCDAALVTEELHLITFETE--VYHOGKIDL 453
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 396 MEESNNGSLSAEFKHLTREORCGNGRANCDAALVTEELHLITFETE--VYHOGKIDL 453

OY 405 YHQ-RTGTLNHFRRMSLKRKIK-----RADRRGASVTEEKFTVLFEQSFVSGNELVFO 458
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 405 YHQ-RTGTLNHFRRMSLKRKIK-----RADRRGASVTEEKFTVLFEQSFVSGNELVFO 458

OY 454 LETHSLPVPVVISNICOMPANASILMYMLTNKNVNFTRKPRIGTDQYAEVLISQFSS 513
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 454 LETHSLPVPVVISNICOMPANASILMYMLTNKNVNFTRKPRIGTDQYAEVLISQFSS 513

OY 459 VKTSLPVPVIVHGSQDNNAATATVLMDNAFA-EPGRVP-FAVPDKVLMPOLCEALNMKEK 516
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 459 VKTSLPVPVIVHGSQDNNAATATVLMDNAFA-EPGRVP-FAVPDKVLMPOLCEALNMKEK 516

OY 514 S--TTKRC SIEQTLTTLAEKLLGPG-----VNSGCOITMAKFCENMAGKGFSPWVLND 567
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 514 S--TTKRC SIEQTLTTLAEKLLGPG-----VNSGCOITMAKFCENMAGKGFSPWVLND 567

OY 517 AEVOSNRGLTKENLFLAOKLFPNSSSHLEDYNSMSVMSQFNRENLPGMNTFTWQMFQGV 576
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 517 AEVOSNRGLTKENLFLAOKLFPNSSSHLEDYNSMSVMSQFNRENLPGMNTFTWQMFQGV 576

OY 568 IIDLVKKTALANEGYIMGFTSKERERAILSTKPRGTFLLRFSSSEKGVTFWTVVERKDI 627
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 568 IIDLVKKTALANEGYIMGFTSKERERAILSTKPRGTFLLRFSSSEKGVTFWTVVERKDI 627

OY 577 VMEVLKHKHRRHMNDGAILGFVNKQOAHDLINKPDCTFLFRSD-SEIGGITIAW--KF 633
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 577 VMEVLKHKHRRHMNDGAILGFVNKQOAHDLINKPDCTFLFRSD-SEIGGITIAW--KF 633

OY 628 ISGKTQIOSVEPYTKOOLNNMSFAELIMGYKIMDATNILVSPLYLYDPIDKEEAFGKYC 687
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 628 ISGKTQIOSVEPYTKOOLNNMSFAELIMGYKIMDATNILVSPLYLYDPIDKEEAFGKYC 687

OY 634 DSPDRNLMLNPKFTTRDPSISLAD-----RLGD-----LNTLIVFPDRKDEVFESKY 683
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 634 DSPDRNLMLNPKFTTRDPSISLAD-----RLGD-----LNTLIVFPDRKDEVFESKY 683

OY 688 RPE-----SQEHPE-----ADPGSAAPYLKTKFCVTPPTCSNTIDLPMSPT 728
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 688 RPE-----SQEHPE-----ADPGSAAPYLKTKFCVTPPTCSNTIDLPMSPT 728

OY 684 TPVLAKAVDGVKPKQKVPEFVNASTDAGASATYMDQ---APSPVACQPQ-HYNNMPPN 739
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 684 TPVLAKAVDGVKPKQKVPEFVNASTDAGASATYMDQ---APSPVACQPQ-HYNNMPPN 739

OY 729 RTDSLMOFGNCGAEPASAGOFESLTFPMDLTSECA 766
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 740 QNPDPVLD--QDGE-----FDLDETMVYA 761

RESULT 10
I57557
DNA-Binding Protein and transcription factor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 28-Jul-2000
C:Accession: I57557
R:Quellle, F.W.; Shimoda, K.; Thieffelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; CI
Mol. Cell. Biol. 15, 3336-3343, 1995
A:Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine pho
A:Reference number: I57557; MUID:95280934
A:Accession: I57557
A>Status: preliminary; translated from GB/EMBL/DBJ

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 14:30:33 ; Search time 23.01 Seconds
(without alignments)
753,044 Million cell updates/sec

Title: US-08-212-185-12

Perfect score: 4029
Sequence: 1 MAOMNOLQDLTRYLKQHLQ.....QFSLTFMDLTSCATSPM 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCFUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4029	100.0	770	1	US-08-369-796-12 Sequence 12, Appl
2	4029	100.0	770	2	US-08-852-091-12 Sequence 12, Appl
3	4029	100.0	770	3	US-08-820-754-12 Sequence 12, Appl
4	4029	100.0	770	3	US-08-956-652-12 Sequence 12, Appl
5	4029	100.0	770	3	US-08-956-869-12 Sequence 12, Appl
6	4029	100.0	770	3	US-08-948-547-12 Sequence 12, Appl
7	4029	100.0	770	4	US-09-364-970-12 Sequence 3, Appl
8	4029	100.0	770	5	PCT-US95-17025-12 Sequence 12, Appl
9	4018	99.7	770	3	US-09-012-710-8 Sequence 8, Appl
10	4016	99.4	770	4	US-09-364-970-5 Sequence 9, Appl
11	4006	99.4	770	1	US-08-416-5818-9 Sequence 5, Appl
12	3977	98.7	770	1	US-08-416-5818-1 Sequence 1, Appl
13	3977	98.7	770	1	US-08-416-5818-5 Sequence 5, Appl
14	3977	98.7	770	4	US-09-087-465-6 Sequence 6, Appl
15	3977	98.7	771	1	US-08-276-099A-14 Sequence 14, Appl
16	3977	98.7	771	1	US-08-781-890-14 Sequence 4, Appl
17	2017	50.1	750	2	US-08-369-796-4 Sequence 4, Appl
18	2017	50.1	750	2	US-08-852-091-4 Sequence 4, Appl
19	2017	50.1	750	2	US-08-820-754-4 Sequence 4, Appl
20	2017	50.1	750	3	US-08-956-652-4 Sequence 4, Appl
21	2017	50.1	750	3	US-08-956-869-4 Sequence 4, Appl
22	2017	50.1	750	3	US-08-948-547-4 Sequence 4, Appl
23	2017	50.1	750	3	US-09-087-465-2 Sequence 2, Appl
24	2017	50.1	750	4	US-09-364-970-1 Sequence 4, Appl
25	2017	50.1	750	5	PCT-US95-17025-4 Sequence 8, Appl
26	2004	49.7	750	4	US-09-364-970-8 Sequence 8, Appl
27	1997	49.6	712	1	US-08-369-796-6 Sequence 6, Appl

28	1997	49.6	712	2	US-08-852-091-6 Sequence 6, Appl
29	1997	49.6	712	2	US-08-820-754-6 Sequence 6, Appl
30	1997	49.6	712	3	US-08-956-652-6 Sequence 6, Appl
31	1997	49.6	712	3	US-08-956-869-6 Sequence 6, Appl
32	1997	49.6	712	3	US-08-948-547-6 Sequence 6, Appl
33	1997	49.6	712	5	PCT-US95-17025-6 Sequence 6, Appl
34	1984	49.2	749	1	US-08-369-796-8 Sequence 8, Appl
35	1984	49.2	749	2	US-08-852-091-8 Sequence 8, Appl
36	1984	49.2	749	3	US-08-820-754-8 Sequence 8, Appl
37	1984	49.2	749	3	US-08-956-652-8 Sequence 8, Appl
38	1984	49.2	749	3	US-08-956-869-8 Sequence 8, Appl
39	1984	49.2	749	3	US-09-012-710-7 Sequence 7, Appl
40	1984	49.2	749	3	US-08-948-547-8 Sequence 8, Appl
41	1984	49.2	749	5	PCT-US95-17025-8 Sequence 8, Appl
42	1971.5	48.9	740	1	US-08-276-099A-12 Sequence 12, Appl
43	1971.5	48.9	740	1	US-08-781-890-12 Sequence 2, Appl
44	1798	44.6	748	1	US-08-408-318-2 Sequence 2, Appl
45	1798	44.6	748	1	US-08-839-164-2 Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-08-369-796-12
: Sequence 12, Application US/08369796
: Patent No. 5716622
: GENERAL INFORMATION:
: APPLICANT: James E. Darnell, Jr.
: APPLICANT: Zilong Wen
: APPLICANT: Curt M. Horvath
: APPLICANT: Zhong Zhong
: TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
: TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
: NUMBER OF SEQUENCES: 39
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.0
: CURRENT PATENT DATA:
: APPLIC NUMBER: US/08/369,796
: FILING NUMBER: 06-JAN-1995
: CLASSIFICATION: 4: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE: PCT NUMBER: 600-1-116
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
: TELETYPE:
: INFORMATION: SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: coding acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-369-796-12
```

Query Match 100.0%; Score 4029; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Caps 0;

```
|||||
Db 1 MAOMNQLOQLDTRRYLKQHLQYSDTFPMEILKQFLAPWIESQDMVAASKESHATLVFHNL 60
OY 61 LGEIDDOYSRFLQESNVLYOHNLRRRIKQFLOSRYLEKPMETARIYARCLWESRLQTA 120
Db 61 LGEIDDOYSRFLQESNVLYOHNLRRRIKQFLOSRYLEKPMETARIYARCLWESRLQTA 120
OY 121 TAAOOGGQANHTAAYVTEKQOMLEQHLQDYKRVODLEOKMKVVENLODFDFNYKTLK 180
Db 121 TAAOOGGQANHTAAYVTEKQOMLEQHLQDYKRVODLEOKMKVVENLODFDFNYKTLK 180
OY 181 SOGDMQDLNGNNOYSTROKMOQLEQMLTALDOMRRSIVSELAGLSAMEYVOKTLTDEEL 240
Db 181 SOGDMQDLNGNNOYSTROKMOQLEQMLTALDOMRRSIVSELAGLSAMEYVOKTLTDEEL 240
OY 241 ADMKRRPEIACIGCPNCLDRLENNITSLAESOLQTRQOIKKLELOOKVSYKGDPIVQ 300
Db 241 ADMKRRPEIACIGCPNCLDRLENNITSLAESOLQTRQOIKKLELOOKVSYKGDPIVQ 300
OY 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVKTGVQFTTKVRLVFPPELNY 360
Db 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVKTGVQFTTKVRLVFPPELNY 360
OY 361 OLKIKVCIKDSGDVAALRGSRKFNILGTNTKVMMEESNNGSLSAEFKHLTLREORCGN 420
Db 361 OLKIKVCIKDSGDVAALRGSRKFNILGTNTKVMMEESNNGSLSAEFKHLTLREORCGN 420
OY 421 GGRACDASLIYTEELHLITFETEVYHOGKIDLETSHLPVVVISNCOMPMASILMY 480
Db 421 GGRACDASLIYTEELHLITFETEVYHOGKIDLETSHLPVVVISNCOMPMASILMY 480
OY 481 NMLTNPNKVNNEFTKPPIGTMDOVAEVLWSQFSSTTKRGLSIEQJTLTAEKLLGCVNYS 540
Db 481 NMLTNPNKVNNEFTKPPIGTMDOVAEVLWSQFSSTTKRGLSIEQJTLTAEKLLGCVNYS 540
OY 541 GCQITWAFCKENMAGKGFSPVWMLDNIIDLVKYIILALMNGYIMGFIISKREBALIST 600
Db 541 GCQITWAFCKENMAGKGFSPVWMLDNIIDLVKYIILALMNGYIMGFIISKREBALIST 600
OY 601 KPPGTFLLRFSESSKEGVTFTWVEKDLSGKTQIOSVEPYTKQOLNNNSFAEIIINGYKIM 660
Db 601 KPPGTFLLRFSESSKEGVTFTWVEKDLSGKTQIOSVEPYTKQOLNNNSFAEIIINGYKIM 660
OY 661 DATNILVSPLYLYPDIPKEAFKCYCRPESOEHPADPGSAAPYLKTKFCVTPPTCSN 720
Db 661 DATNILVSPLYLYPDIPKEAFKCYCRPESOEHPADPGSAAPYLKTKFCVTPPTCSN 720
OY 721 TIDLPMSPRTLDLSLMQFNGNGGAPPSAGGOFESLTFPMDLTSECATSPM 770
Db 721 TIDLPMSPRTLDLSLMQFNGNGGAPPSAGGOFESLTFPMDLTSECATSPM 770

RESULT 2
US-08-852-091-12
Sequence 12 Application US/08852091
Patent No. 5683228
GENERAL INFORMATION:
APPLICANT: James E. Darnell, Jr.
APPLICANT: Zilong Wen
APPLICANT: Curt M. Horvath
APPLICANT: Zhong Zhong
TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESS: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,091
FILING DATE: 06-MAY-1997
CLASSIFICATION: 424
PRIORITY INFORMATION:
APPLICATION NUMBER: 08/369,796
FILING DATE: 06-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-116
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-852-091-12
```

```
Query Match 100.0%; Score 4029; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAOMNQLOQLDTRRYLKQHLQYSDTFPMEILKQFLAPWIESQDMVAASKESHATLVFHNL 60
Db 1 MAOMNQLOQLDTRRYLKQHLQYSDTFPMEILKQFLAPWIESQDMVAASKESHATLVFHNL 60
OY 61 LGEIDDOYSRFLQESNVLYOHNLRRRIKQFLOSRYLEKPMETARIYARCLWESRLQTA 120
Db 61 LGEIDDOYSRFLQESNVLYOHNLRRRIKQFLOSRYLEKPMETARIYARCLWESRLQTA 120
OY 121 TAAOOGGQANHTAAYVTEKQOMLEQHLQDYKRVODLEOKMKVVENLODFDFNYKTLK 180
Db 121 TAAOOGGQANHTAAYVTEKQOMLEQHLQDYKRVODLEOKMKVVENLODFDFNYKTLK 180
OY 181 SOGDMQDLNGNNOYSTROKMOQLEQMLTALDOMRRSIVSELAGLSAMEYVOKTLTDEEL 240
Db 181 SOGDMQDLNGNNOYSTROKMOQLEQMLTALDOMRRSIVSELAGLSAMEYVOKTLTDEEL 240
OY 241 ADMKRRPEIACIGCPNCLDRLENNITSLAESOLQTRQOIKKLELOOKVSYKGDPIVQ 300
Db 241 ADMKRRPEIACIGCPNCLDRLENNITSLAESOLQTRQOIKKLELOOKVSYKGDPIVQ 300
OY 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVKTGVQFTTKVRLVFPPELNY 360
Db 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVKTGVQFTTKVRLVFPPELNY 360
OY 361 OLKIKVCIKDSGDVAALRGSRKFNILGTNTKVMMEESNNGSLSAEFKHLTLREORCGN 420
Db 361 OLKIKVCIKDSGDVAALRGSRKFNILGTNTKVMMEESNNGSLSAEFKHLTLREORCGN 420
OY 421 GGRACDASLIYTEELHLITFETEVYHOGKIDLETSHLPVVVISNCOMPMASILMY 480
Db 421 GGRACDASLIYTEELHLITFETEVYHOGKIDLETSHLPVVVISNCOMPMASILMY 480
OY 481 NMLTNPNKVNNEFTKPPIGTMDOVAEVLWSQFSSTTKRGLSIEQJTLTAEKLLGCVNYS 540
Db 481 NMLTNPNKVNNEFTKPPIGTMDOVAEVLWSQFSSTTKRGLSIEQJTLTAEKLLGCVNYS 540
OY 541 GCQITWAFCKENMAGKGFSPVWMLDNIIDLVKYIILALMNGYIMGFIISKREBALIST 600
Db 541 GCQITWAFCKENMAGKGFSPVWMLDNIIDLVKYIILALMNGYIMGFIISKREBALIST 600
OY 601 KPPGTFLLRFSESSKEGVTFTWVEKDLSGKTQIOSVEPYTKQOLNNNSFAEIIINGYKIM 660
Db 601 KPPGTFLLRFSESSKEGVTFTWVEKDLSGKTQIOSVEPYTKQOLNNNSFAEIIINGYKIM 660
```

```
DB 601 KPPGFTLRLRSESSKEGCVFTTWKEDISGKTQIOSVEPYTKOQLNNMSFAELIMGYKIM 660
OY 661 DATNIIIVSLVLYLPDIPIKEAFGKYCRPESQEHPEADGSAAPYLKTKFCVPTTCSN 720
DB 661 DATNIIIVSLVLYLPDIPIKEAFGKYCRPESQEHPEADGSAAPYLKTKFCVPTTCSN 720
OY 721 TIDLPMSPPTLDSLMQFGNNGEAGPSAGGOFESLTFDMDLTSECATSPM 770
DB 721 TIDLPMSPPTLDSLMQFGNNGEAGPSAGGOFESLTFDMDLTSECATSPM 770

RESULT
US-08-820-754-12
; Sequence 12, Application US/08820754
; Patent No. 597835
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,754
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,185
; FILING DATE: 11-MAR-1994
; APPLICATION NUMBER: US 07/980,498
; FILING DATE: 23-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; APPLICATION DATA:
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-820-754-12

Query Match 100.0%; Score 4029; DB 2: Length 770;
Best Local Similarity 100.0%; Pred. No. 0;
```

```
Matches 770: insertive 0: mismatches 0: indels 0: gaps 0:
OY 1 MAQMNQLOOLDTRYLYKOLHOLYSDIFPMELROFLAPWIESQDMAYAAKESHATLVFHNL 60
DB 1 MAQMNQLOOLDTRYLYKOLHOLYSDIFPMELROFLAPWIESQDMAYAAKESHATLVFHNL 60
OY 61 LGEIDQOYSRFLQESNVLYOHNLRRIKOFLORSYLEKPEIARIRARCLMESRLQTPAA 120
DB 61 LGEIDQOYSRFLQESNVLYOHNLRRIKOFLORSYLEKPEIARIRARCLMESRLQTPAA 120
OY 121 TAAOQGGQANHPPTAAVYTEKQOMLEOHLODVRKRYODLKKKKVVENLODDDFNFKTLK 180
DB 121 TAAOQGGQANHPPTAAVYTEKQOMLEOHLODVRKRYODLKKKKVVENLODDDFNFKTLK 180
OY 181 SOGDMDLGNNOVSVRKMOOLEOMLTALDOMKRSIVSELAGLISAMEYVOKTLTDEEL 240
DB 181 SOGDMDLGNNOVSVRKMOOLEOMLTALDOMKRSIVSELAGLISAMEYVOKTLTDEEL 240
OY 241 ADMKRRPEIACIGPPNICLDRLENMITSLESOLOTRQOIKKLELOQVSKGDPITVO 300
DB 241 ADMKRRPEIACIGPPNICLDRLENMITSLESOLOTRQOIKKLELOQVSKGDPITVO 300
OY 301 HRPMLFERIVLEFRNLKSAFVEROPCPMHHPDRPLVIKTSVOFTTKYRLLVKPEPELV 360
DB 301 HRPMLFERIVLEFRNLKSAFVEROPCPMHHPDRPLVIKTSVOFTTKYRLLVKPEPELV 360
OY 361 OLKIKVCIDKSDGVVALRGSRRFNLTGNTVKMNFESNNOSLSAFPHLTLRORCGN 420
DB 361 OLKIKVCIDKSDGVVALRGSRRFNLTGNTVKMNFESNNOSLSAFPHLTLRORCGN 420
OY 421 GGRANDASLIVTEELHLITFETEVYHOGIKIDLETHSIPVVVISNCOMPANASTLMY 480
DB 421 GGRANDASLIVTEELHLITFETEVYHOGIKIDLETHSIPVVVISNCOMPANASTLMY 480
OY 481 NMILTNPKNVNPFTKPPITGTDQVAEVLISWQSSFTTKRGLSTHOUTTLAEKILGCVNVS 540
DB 481 NMILTNPKNVNPFTKPPITGTDQVAEVLISWQSSFTTKRGLSTHOUTTLAEKILGCVNVS 540
OY 541 GCOITWAKFCENMAGKGSFWWMLDNIIDLVKKYIILAMNGCYIMGFISKREBAIIST 600
DB 541 GCOITWAKFCENMAGKGSFWWMLDNIIDLVKKYIILAMNGCYIMGFISKREBAIIST 600
OY 601 KPPGFTLRLRSESSKEGCVFTTWKEDISGKTQIOSVEPYTKOQLNNMSFAELIMGYKIM 660
DB 601 KPPGFTLRLRSESSKEGCVFTTWKEDISGKTQIOSVEPYTKOQLNNMSFAELIMGYKIM 660
OY 661 DATNIIIVSLVLYLPDIPIKEAFGKYCRPESQEHPEADGSAAPYLKTKFCVPTTCSN 720
DB 661 DATNIIIVSLVLYLPDIPIKEAFGKYCRPESQEHPEADGSAAPYLKTKFCVPTTCSN 720
OY 721 TIDLPMSPPTLDSLMQFGNNGEAGPSAGGOFESLTFDMDLTSECATSPM 770
DB 721 TIDLPMSPPTLDSLMQFGNNGEAGPSAGGOFESLTFDMDLTSECATSPM 770

RESULT
US-08-956-652-12
; Sequence 12, Application US/08956652
; Patent No. 6013475
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
```

COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,652
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,185
FILING DATE: 11-MAR-1994
APPLICATION NUMBER: US 07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-956-652-12

Query Match 100.0%; Score 4029; DB 3; length 770;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAOMNLOOLDTRTYLKQHLNLYSDTFPMELROFLAPMIESQDMAAASKESNATLVFHNL 60
DB 1 MAOMNLOOLDTRTYLKQHLNLYSDTFPMELROFLAPMIESQDMAAASKESNATLVFHNL 60
QY 61 LGEIDDOYSEFLDOESNVLVQHNLRRIKOFLOSRYLEKPMELIARIVARCLMEESRLLOTAA 120
DB 61 LGEIDDOYSEFLDOESNVLVQHNLRRIKOFLOSRYLEKPMELIARIVARCLMEESRLLOTAA 120
QY 121 TAAOGGGAANHPTAAVUTEKQMLECHLQDVRRVODLEQKMKVVENLQDDPFNTKTLK 180
DB 121 TAAOGGGAANHPTAAVUTEKQMLECHLQDVRRVODLEQKMKVVENLQDDPFNTKTLK 180
QY 181 SOGDMDLNGNNSVTRQKMOOLEOMLTALDOMRSIVSLAGLSAMEYVOKTLTDEEL 240
DB 181 SOGDMDLNGNNSVTRQKMOOLEOMLTALDOMRSIVSLAGLSAMEYVOKTLTDEEL 240
QY 241 ADMKRRPEIACIGSPNICDLRLNMTTSLAESOLOTRQOIKKLEELQOKVSYKGPYVQ 300
DB 241 ADMKRRPEIACIGSPNICDLRLNMTTSLAESOLOTRQOIKKLEELQOKVSYKGPYVQ 300
QY 301 HRPMLEERIVELFRNLKMSAFVVERQPCMHNDPRPLVITGVQFTTKVRLVKFPPLNY 360
DB 301 HRPMLEERIVELFRNLKMSAFVVERQPCMHNDPRPLVITGVQFTTKVRLVKFPPLNY 360
QY 361 QLKIKVCIDKSDGVVALRSRKFNIIGTNTKYNMMEESNNGSLSAEFKLLTLREORCGN 420
DB 361 QLKIKVCIDKSDGVVALRSRKFNIIGTNTKYNMMEESNNGSLSAEFKLLTLREORCGN 420

QY 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNCOMPMANASILMY 480
DB 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNCOMPMANASILMY 480
QY 481 NMLTNPKVNVNFTKPPIGTMOQVAEVLMSOPSSITKRLSTLEOLTTLAEKLLGCVNYS 540
DB 481 NMLTNPKVNVNFTKPPIGTMOQVAEVLMSOPSSITKRLSTLEOLTTLAEKLLGCVNYS 540
QY 541 GCQITWAKFCKENMAGKGFSPVWLDNIIDLVKKYILALMNGYIMGFLSKRERAILST 600
DB 541 GCQITWAKFCKENMAGKGFSPVWLDNIIDLVKKYILALMNGYIMGFLSKRERAILST 600
QY 601 KPPGFTLLRFSSSKEGVYTFWVERKDISGKTQIOSVEPYTQOULNNSFAELIMGYKIM 660
DB 601 KPPGFTLLRFSSSKEGVYTFWVERKDISGKTQIOSVEPYTQOULNNSFAELIMGYKIM 660
QY 661 DATNILVSPLVLYLPDIPKEAFKCYCRPESQHPHADGSAAPYLKTFICVPTTCSN 720
DB 661 DATNILVSPLVLYLPDIPKEAFKCYCRPESQHPHADGSAAPYLKTFICVPTTCSN 720
QY 721 TIDLPMSPRTLDLSMOFGNNGGAEPSSAGOPESLTFMDLTSECATSPM 770
DB 721 TIDLPMSPRTLDLSMOFGNNGGAEPSSAGOPESLTFMDLTSECATSPM 770

RESULT 5
US-08-956-869-12
Sequence 12, Application US/08956869
Patent No. 6030808
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,869
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/212,185
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-956-869-12

Query Match 100.0%; Score 4029; DB 3; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MAONNOLOOUDTRYKOLHOLYSDFPMLROFLAPWIESODMAYAAKESHATLVFNNL 60
DB 1 MAONNOLOOUDTRYKOLHOLYSDFPMLROFLAPWIESODMAYAAKESHATLVFNNL 60
OY 61 LGEIDOOVSRLFOESNVLYOHNLRRIKOFLORSYLEKPMELIARIYARCLMEESRLLOTAA 120
DB 61 LGEIDOOVSRLFOESNVLYOHNLRRIKOFLORSYLEKPMELIARIYARCLMEESRLLOTAA 120
OY 121 TAAOOGGQANHPTAAVYTEKQOMLEOHLODYRKRYODLEOKMKVVENLODDDFNYKTLK 180
DB 121 TAAOOGGQANHPTAAVYTEKQOMLEOHLODYRKRYODLEOKMKVVENLODDDFNYKTLK 180
OY 181 SOGDMODLNGNOSVTRKMOOLEOMLTALDOMRKSIVSELAGLLSAMERYOKTLTDEEL 240
DB 181 SOGDMODLNGNOSVTRKMOOLEOMLTALDOMRKSIVSELAGLLSAMERYOKTLTDEEL 240
OY 241 ADMKRREPIACIGPPNLCIDRLLENMITSLSAESOLQTRQOLKLELOOKSYKCDPIVQ 300
DB 241 ADMKRREPIACIGPPNLCIDRLLENMITSLSAESOLQTRQOLKLELOOKSYKCDPIVQ 300
OY 301 HRPMLEERIVELFRNLKMSAFVEROPCPMPHPRPLVKTGOVFTTKVRLVLPPELNY 360
DB 301 HRPMLEERIVELFRNLKMSAFVEROPCPMPHPRPLVKTGOVFTTKVRLVLPPELNY 360
OY 361 OLKIKVCIDKSGVAAALRGSRKFENLGTNTKVMMEESNNGSLSAEKKHLTLRPORCGN 420
DB 361 OLKIKVCIDKSGVAAALRGSRKFENLGTNTKVMMEESNNGSLSAEKKHLTLRPORCGN 420
OY 421 GGRANCASLIVTEELHLITFEFEVYHOGKLTIDETHSLPVVVISNICOMPMAASILMY 480
DB 421 GGRANCASLIVTEELHLITFEFEVYHOGKLTIDETHSLPVVVISNICOMPMAASILMY 480
OY 481 NMLTNNRNKRVNFTKPPITGTDQVAEVLSPQSSSTTKRGCLSTEOLTTAEKLLGCVNYS 540
DB 481 NMLTNNRNKRVNFTKPPITGTDQVAEVLSPQSSSTTKRGCLSTEOLTTAEKLLGCVNYS 540
OY 541 GCOTITWAKFCCKENNAKGKFSFWMLNDITIDLVKTYLLALMEGVIMGFTISKREPAILST 600
DB 541 GCOTITWAKFCCKENNAKGKFSFWMLNDITIDLVKTYLLALMEGVIMGFTISKREPAILST 600
OY 601 KPPGCTFLFRFSESSKEGCVFTTWEEKDISGKTQIOSVEPYTKQOLNNNSFAEIIIMGYKIM 660
DB 601 KPPGCTFLFRFSESSKEGCVFTTWEEKDISGKTQIOSVEPYTKQOLNNNSFAEIIIMGYKIM 660
OY 661 DATNIVLSPLVYLPDIKPEAFKCYKCRPESEOEHPADPGSAAPLAKTKFCVPTTCSN 720
DB 661 DATNIVLSPLVYLPDIKPEAFKCYKCRPESEOEHPADPGSAAPLAKTKFCVPTTCSN 720
OY 721 TIDLPMSRITLDSLMQFGNNGEGAPPSAGOFESLTFMDLTSECATSPM 770
DB 721 TIDLPMSRITLDSLMQFGNNGEGAPPSAGOFESLTFMDLTSECATSPM 770

```

RESULT 6
 US-08-948-547-12
 : Sequence 12, Application us/08948547
 : Patent No. 6124118
 : GENERAL INFORMATION:

```

APPLICANT: Arnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Men, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
NUMBER OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauder & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,547
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/212,185
FILING DATE: 11-MAR-1994
APPLICATION NUMBER: US/07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO/US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/MARKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
INFORMATION:
SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-948-547-12

```

Query Match 100.0%; Score 4029; DB 3; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MAONNOLOOUDTRYKOLHOLYSDFPMLROFLAPWIESODMAYAAKESHATLVFNNL 60
DB 1 MAONNOLOOUDTRYKOLHOLYSDFPMLROFLAPWIESODMAYAAKESHATLVFNNL 60
OY 61 LGEIDOOVSRLFOESNVLYOHNLRRIKOFLORSYLEKPMELIARIYARCLMEESRLLOTAA 120
DB 61 LGEIDOOVSRLFOESNVLYOHNLRRIKOFLORSYLEKPMELIARIYARCLMEESRLLOTAA 120
OY 121 TAAOOGGQANHPTAAVYTEKQOMLEOHLODYRKRYODLEOKMKVVENLODDDFNYKTLK 180
DB 121 TAAOOGGQANHPTAAVYTEKQOMLEOHLODYRKRYODLEOKMKVVENLODDDFNYKTLK 180
OY 181 SOGDMODLNGNOSVTRKMOOLEOMLTALDOMRKSIVSELAGLLSAMERYOKTLTDEEL 240
DB 181 SOGDMODLNGNOSVTRKMOOLEOMLTALDOMRKSIVSELAGLLSAMERYOKTLTDEEL 240

```

Db 181 SOGDMODLNGNNSVTRKMOOLEOMLTALDOMRSTVSELAGLSAMEYOKLTDEEL 240
QY 241 ADMKRRPEIACIGGPPNCLDRLENMWITSLAESOLQTRQIJKLEELQOKVSKGDIYQ 300
Db 241 ADMKRRPEIACIGGPPNCLDRLENMWITSLAESOLQTRQIJKLEELQOKVSKGDIYQ 300
QY 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVKTGVQFTTKVRLVKEPELNY 360
Db 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVKTGVQFTTKVRLVKEPELNY 360
QY 361 QLKIKVCIDKDSGDVAALRGSRKFNILGTNTKVMNMEESNNGSLSAEKKHLTLREORCGN 420
Db 361 QLKIKVCIDKDSGDVAALRGSRKFNILGTNTKVMNMEESNNGSLSAEKKHLTLREORCGN 420
QY 421 GGRANCASLIYTELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPMAASILMY 480
Db 421 GGRANCASLIYTELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPMAASILMY 480
QY 481 NMLTNNPKNVNFTKPPIGTWDQVAEVLSPQSSSTTKRGLSIEOLTTLAEKLLGPGVNS 540
Db 481 NMLTNNPKNVNFTKPPIGTWDQVAEVLSPQSSSTTKRGLSIEOLTTLAEKLLGPGVNS 540
QY 541 GCQITWAFCKENMAGKGFSPWVWLDNIIDLVKKYILALMNEGYIMGFISKERERAILST 600
Db 541 GCQITWAFCKENMAGKGFSPWVWLDNIIDLVKKYILALMNEGYIMGFISKERERAILST 600
QY 601 KPPGTFLLRFSESSKEGGVFTFWKEKDISGKTQIOSVEPYTKOOLNNSPAETIMGYKIM 660
Db 601 KPPGTFLLRFSESSKEGGVFTFWKEKDISGKTQIOSVEPYTKOOLNNSPAETIMGYKIM 660
QY 661 DATNLVSPVLYLPDIDKEEAFKCYCRPESOEHPDPSAAPYLKTKFCVPTTCSN 720
Db 661 DATNLVSPVLYLPDIDKEEAFKCYCRPESOEHPDPSAAPYLKTKFCVPTTCSN 720
QY 721 TIDLPMSPRTLDLSLMQFGNNGEGAPPSAGGFESLTFDMDLTSECATSPM 770
Db 721 TIDLPMSPRTLDLSLMQFGNNGEGAPPSAGGFESLTFDMDLTSECATSPM 770
RESULT 7
US-09-364-970-3
: Sequence 3, Application US/09364970
: Patent No. 6235873
: GENERAL INFORMATION:
: APPLICANT: Bromberg, Jacqueline
: TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
: TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
: FILE REFERENCE: 600-1-252
: CURRENT APPLICATION NUMBER: US/09/364,970
: CURRENT FILING DATE: 1999-07-31
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 770
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-364-970-3
Query Match 100.0%: Score 4029; DB 4: Length 770:
Best Local Similarity 100.0%: Pred. No. 0:
Matches 770: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 121 TAAOOGGQANHPTAAVYTEKQOMLEOHLODYRKRVODLEQKMKVVENLODDDFPNKTLK 180
QY 181 SOGDMODLNGNNSVTRKMOOLEOMLTALDOMRSTVSELAGLSAMEYOKLTDEEL 240
Db 181 SOGDMODLNGNNSVTRKMOOLEOMLTALDOMRSTVSELAGLSAMEYOKLTDEEL 240
QY 241 ADMKRRPEIACIGGPPNCLDRLENMWITSLAESOLQTRQIJKLEELQOKVSKGDIYQ 300
Db 241 ADMKRRPEIACIGGPPNCLDRLENMWITSLAESOLQTRQIJKLEELQOKVSKGDIYQ 300
QY 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVKTGVQFTTKVRLVKEPELNY 360
Db 301 HRPMLEERIVELFRNLKMSAFVEROPCPMHPDRPLVKTGVQFTTKVRLVKEPELNY 360
QY 361 QLKIKVCIDKDSGDVAALRGSRKFNILGTNTKVMNMEESNNGSLSAEKKHLTLREORCGN 420
Db 361 QLKIKVCIDKDSGDVAALRGSRKFNILGTNTKVMNMEESNNGSLSAEKKHLTLREORCGN 420
QY 421 GGRANCASLIYTELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPMAASILMY 480
Db 421 GGRANCASLIYTELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPMAASILMY 480
QY 481 NMLTNNPKNVNFTKPPIGTWDQVAEVLSPQSSSTTKRGLSIEOLTTLAEKLLGPGVNS 540
Db 481 NMLTNNPKNVNFTKPPIGTWDQVAEVLSPQSSSTTKRGLSIEOLTTLAEKLLGPGVNS 540
QY 541 GCQITWAFCKENMAGKGFSPWVWLDNIIDLVKKYILALMNEGYIMGFISKERERAILST 600
Db 541 GCQITWAFCKENMAGKGFSPWVWLDNIIDLVKKYILALMNEGYIMGFISKERERAILST 600
QY 601 KPPGTFLLRFSESSKEGGVFTFWKEKDISGKTQIOSVEPYTKOOLNNSPAETIMGYKIM 660
Db 601 KPPGTFLLRFSESSKEGGVFTFWKEKDISGKTQIOSVEPYTKOOLNNSPAETIMGYKIM 660
QY 661 DATNLVSPVLYLPDIDKEEAFKCYCRPESOEHPDPSAAPYLKTKFCVPTTCSN 720
Db 661 DATNLVSPVLYLPDIDKEEAFKCYCRPESOEHPDPSAAPYLKTKFCVPTTCSN 720
QY 721 TIDLPMSPRTLDLSLMQFGNNGEGAPPSAGGFESLTFDMDLTSECATSPM 770
Db 721 TIDLPMSPRTLDLSLMQFGNNGEGAPPSAGGFESLTFDMDLTSECATSPM 770
RESULT 8
PCT-US95-17025-12
: Sequence 12, Application PC/TUS9517025
: GENERAL INFORMATION:
: APPLICANT: James E. Darnell, Jr.
: APPLICANT: Zilong Wen
: APPLICANT: Curt M. Horvath
: APPLICANT: Zhong Zhong
: TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
: TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
: NUMBER OF SEQUENCES: 39
: CORRESPONDENT ADDRESS:
: ADDRESS: Klauder & Jackson
: STREET: 601 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER RELEASABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/17025
: FILING DATE: 28-DEC-1995
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/369,796

```

: FILING DATE: 06-JAN-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-116
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
: TELEX: 133521
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-17025-12

```

Query Match 100.0%; Score 4029; DB 5; Length 770;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 770: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 MAOMNOLQDLDTTRYKQLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
Db 1 MAOMNOLQDLDTTRYKQLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
Oy 61 LGEIDDOYSRFLQESNVLYOHNLRIKOFLOSRYLEKPEIARIYARCLMEESRLLOTAA 120
Db 61 LGEIDDOYSRFLQESNVLYOHNLRIKOFLOSRYLEKPEIARIYARCLMEESRLLOTAA 120
Oy 121 TAAOOGGOANHPTAAVVTTEKQOMLEOHLODVRRKRVODLEOKMKVVENLODDDFNFKTLK 180
Db 121 TAAOOGGOANHPTAAVVTTEKQOMLEOHLODVRRKRVODLEOKMKVVENLODDDFNFKTLK 180
Oy 181 SOGDMODLGNNSVTRQKMOOLEOMLTALDOMRISIVSELAGLISAMEYVOKTLTDEEL 240
Db 181 SOGDMODLGNNSVTRQKMOOLEOMLTALDOMRISIVSELAGLISAMEYVOKTLTDEEL 240
Oy 241 ADMKRRPEIACIGPPNICIDRLNNMTSLAESQLOTROOIKLEELQOKVSKGDPITVO 300
Db 241 ADMKRRPEIACIGPPNICIDRLNNMTSLAESQLOTROOIKLEELQOKVSKGDPITVO 300
Oy 301 HRPMLEERIVELFRNLMSAFVVEROPCPMPHDPRLVIKTGVQFTTKYRLLVKPPELAY 360
Db 301 HRPMLEERIVELFRNLMSAFVVEROPCPMPHDPRLVIKTGVQFTTKYRLLVKPPELAY 360
Oy 361 OLKIKYCIDKDSQVAALGSRKFNLTGNTKYANNEESNGSLSAEFKHLTLREORCGN 420
Db 361 OLKIKYCIDKDSQVAALGSRKFNLTGNTKYANNEESNGSLSAEFKHLTLREORCGN 420
Oy 421 GGRANDASLTVEELHLITFEFEVYHOGKIDLETHSLPVVYISITICOMPNAASILAY 480
Db 421 GGRANDASLTVEELHLITFEFEVYHOGKIDLETHSLPVVYISITICOMPNAASILAY 480
Oy 481 NMLTNPKNVNFTKPPICGTWDOVAEVLWSQFSSTTKRGSLIEQLTTLAEKLLGPGVNS 540
Db 481 NMLTNPKNVNFTKPPICGTWDOVAEVLWSQFSSTTKRGSLIEQLTTLAEKLLGPGVNS 540
Oy 541 GCOITNAKCKEMMAKGSFWMWLDNIIDLYKYLALAMNEGYINGFISKEBERALIST 600
Db 541 GCOITNAKCKEMMAKGSFWMWLDNIIDLYKYLALAMNEGYINGFISKEBERALIST 600
Oy 601 KPGTFLLPFSESSKEGGVTFVWEKDISGKTIOSEVEPTKQOLNNMSFAELIMGYKIM 660
Db 601 KPGTFLLPFSESSKEGGVTFVWEKDISGKTIOSEVEPTKQOLNNMSFAELIMGYKIM 660
Oy 661 DATNIIIVSPLVLYPIPIKEAEAGKYCRPSOEHPREADPGSAAPYLKTFICVPTTCSN 720
Db 661 DATNIIIVSPLVLYPIPIKEAEAGKYCRPSOEHPREADPGSAAPYLKTFICVPTTCSN 720
Oy 721 TIDLPMSPRTIOSLMQFNGNGEAPPSAGGOFESLTFPMDLTSECATSPM 770
Db 721 TIDLPMSPRTIOSLMQFNGNGEAPPSAGGOFESLTFPMDLTSECATSPM 770

```

Db 721 TIDLPMSPRTIOSLMQFNGNGEAPPSAGGOFESLTFPMDLTSECATSPM 770

RESULT 9

US-09-012-710-8

Sequence 8, Application US/09012710

Patent No. 6087478

GENERAL INFORMATION:

APPLICANT: Vinkemeier, Uwe

APPLICANT: Moarefi, Ismail

APPLICANT: Darnell, Jr., James E.

APPLICANT: Kuriyan, John

TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESS: Klauber & Jackson

STREET: 411 Hackensack Avenue, 4th floor

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/012,710

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-194

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 770 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEICAL: NO

US-09-012-710-8

Query Match 99.7%; Score 4018; DB 3; Length 770;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 768: Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Oy 1 MAOMNOLQDLDTTRYKQLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
Db 1 MAOMNOLQDLDTTRYKQLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
Oy 61 LGEIDDOYSRFLQESNVLYOHNLRIKOFLOSRYLEKPEIARIYARCLMEESRLLOTAA 120
Db 61 LGEIDDOYSRFLQESNVLYOHNLRIKOFLOSRYLEKPEIARIYARCLMEESRLLOTAA 120
Oy 121 TAAOOGGOANHPTAAVVTTEKQOMLEOHLODVRRKRVODLEOKMKVVENLODDDFNFKTLK 180
Db 121 TAAOOGGOANHPTAAVVTTEKQOMLEOHLODVRRKRVODLEOKMKVVENLODDDFNFKTLK 180
Oy 181 SOGDMODLGNNSVTRQKMOOLEOMLTALDOMRISIVSELAGLISAMEYVOKTLTDEEL 240
Db 181 SOGDMODLGNNSVTRQKMOOLEOMLTALDOMRISIVSELAGLISAMEYVOKTLTDEEL 240
Oy 241 ADMKRRPEIACIGPPNICIDRLNNMTSLAESQLOTROOIKLEELQOKVSKGDPITVO 300
Db 241 ADMKRRPEIACIGPPNICIDRLNNMTSLAESQLOTROOIKLEELQOKVSKGDPITVO 300

```

```

OY 301 HRPMLEERIVELFRLNLSAFVEROPCPMHDRPLVKTGVQFTTKVRLLVKPEPILNY 360
DB 301 HRPMLEERIVELFRLNLSAFVEROPCPMHDRPLVKTGVQFTTKVRLLVKPEPILNY 360
OY 361 QLKIKVCIDKDSGDVAALGSRKFNILGTNTKVMNMEESNNSLSAEFKHLTLRQRCGN 420
DB 361 QLKIKVCIDKDSGDVAALGSRKFNILGTNTKVMNMEESNNSLSAEFKHLTLRQRCGN 420
OY 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANAMASILMY 480
DB 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANAMASILMY 480
OY 481 NMLTNPKNVNFFTKPRIGTMDQVAEVLISWOPSTTKRGLSIEOLTTLAEKLLGPGVNS 540
DB 481 NMLTNPKNVNFFTKPRIGTMDQVAEVLISWOPSTTKRGLSIEOLTTLAEKLLGPGVNS 540
OY 541 GCOITWAKFCCKENMAGKGFSPFWMLDNIIDLKYYILALMNGYIMGFISKERERAILST 600
DB 541 GCOITWAKFCCKENMAGKGFSPFWMLDNIIDLKYYILALMNGYIMGFISKERERAILST 600
OY 601 KPPGFTLLRFSSSKREGVYFTWVEKDISGKTQIOSVEPYTKOOLNNSFAELIMGYKIM 660
DB 601 KPPGFTLLRFSSSKREGVYFTWVEKDISGKTQIOSVEPYTKOOLNNSFAELIMGYKIM 660
OY 661 DATNLIPLVLYLPDIPKEFAFGKYCRPESQEHPEADPGSAAPYLKTKFCVPTTCSN 720
DB 661 DATNLIPLVLYLPDIPKEFAFGKYCRPESQEHPEADPGSAAPYLKTKFCVPTTCSN 720
OY 721 TIDLPMSPRTLDLSLMQFNGNNGGAEPSSAGOFESLTFPMDLTSSECATSPM 770
DB 721 TIDLPMSPRTLDLSLMQFNGNNGGAEPSSAGOFESLTFPMDLTSSECATSPM 770

```

```

RESULT 10
US-09-364-970-5
: Sequence 5, Application US/09364970
: Patent No. 6235873
: GENERAL INFORMATION:
: APPLICANT: Bromberg, Jacqueline
: TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
: TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
: FILE REFERENCE: 600-1-252
: CURRENT APPLICATION NUMBER: US/09/364,970
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 5
: LENGTH: 770
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-364-970-5

```

```

Query Match          99.7%: Score 4016; DB 4; Length 770;
Best Local Similarity 99.7%: Pred. No. 0;
Matches 768; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 MAOMNLOLOLDRYKLOLQOLYSDPRMELROFLAPWISODMAVAASKESHATLVFHNL 60
DB 1 MAOMNLOLOLDRYKLOLQOLYSDPRMELROFLAPWISODMAVAASKESHATLVFHNL 60
OY 61 LGEIDDOYSRFLOESNVLYOHNLRIKOFLOSRYLEKPMELIARIIVARCLMEESRLQTA 120
DB 61 LGEIDDOYSRFLOESNVLYOHNLRIKOFLOSRYLEKPMELIARIIVARCLMEESRLQTA 120
OY 121 TAAOOGGAANHPTAAVYTEKQOMLEOHLODVRRKRVODLEOKMKVVENLDDPFNFKTLK 180
DB 121 TAAOOGGAANHPTAAVYTEKQOMLEOHLODVRRKRVODLEOKMKVVENLDDPFNFKTLK 180
OY 181 SOGDMDLGNNSQSVTRKMOOLEOMLTALDQMRSTVSELAGILSMEVVOQTLLDEEL 240
DB 181 SOGDMDLGNNSQSVTRKMOOLEOMLTALDQMRSTVSELAGILSMEVVOQTLLDEEL 240

```

```

DB 181 SOGDMDLGNNSQSVTRKMOOLEOMLTALDQMRSTVSELAGILSMEVVOQTLLDEEL 240
OY 241 ADMKRRPEIACIGGPNICIDLENNMITSIAESQOTROOIKKEELOOKSYKNDPIVQ 300
DB 241 ADMKRRPEIACIGGPNICIDLENNMITSIAESQOTROOIKKEELOOKSYKNDPIVQ 300
OY 301 HRPMLEERIVELFRLNLSAFVEROPCPMHDRPLVKTGVQFTTKVRLLVKPEPILNY 360
DB 301 HRPMLEERIVELFRLNLSAFVEROPCPMHDRPLVKTGVQFTTKVRLLVKPEPILNY 360
OY 361 QLKIKVCIDKDSGDVAALGSRKFNILGTNTKVMNMEESNNSLSAEFKHLTLRQRCGN 420
DB 361 QLKIKVCIDKDSGDVAALGSRKFNILGTNTKVMNMEESNNSLSAEFKHLTLRQRCGN 420
OY 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANAMASILMY 480
DB 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANAMASILMY 480
OY 481 NMLTNPKNVNFFTKPRIGTMDQVAEVLISWOPSTTKRGLSIEOLTTLAEKLLGPGVNS 540
DB 481 NMLTNPKNVNFFTKPRIGTMDQVAEVLISWOPSTTKRGLSIEOLTTLAEKLLGPGVNS 540
OY 541 GCOITWAKFCCKENMAGKGFSPFWMLDNIIDLKYYILALMNGYIMGFISKERERAILST 600
DB 541 GCOITWAKFCCKENMAGKGFSPFWMLDNIIDLKYYILALMNGYIMGFISKERERAILST 600
OY 601 KPPGFTLLRFSSSKREGVYFTWVEKDISGKTQIOSVEPYTKOOLNNSFAELIMGYKIM 660
DB 601 KPPGFTLLRFSSSKREGVYFTWVEKDISGKTQIOSVEPYTKOOLNNSFAELIMGYKIM 660
OY 661 DATNLIPLVLYLPDIPKEFAFGKYCRPESQEHPEADPGSAAPYLKTKFCVPTTCSN 720
DB 661 DATNLIPLVLYLPDIPKEFAFGKYCRPESQEHPEADPGSAAPYLKTKFCVPTTCSN 720
OY 721 TIDLPMSPRTLDLSLMQFNGNNGGAEPSSAGOFESLTFPMDLTSSECATSPM 770
DB 721 TIDLPMSPRTLDLSLMQFNGNNGGAEPSSAGOFESLTFPMDLTSSECATSPM 770

```

```

RESULT 11
US-08-416-581B-9
: Sequence 9, Application US/08416581B
: Patent No. 5719042
: GENERAL INFORMATION:
: APPLICANT: Kishimoto, Tadamiatsu
: TITLE OF INVENTION: TRANSCRIPTION FACTOR APRF
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SUGHRUE, MIOW, ZINN, MACPEAK & SEAS
: STREET: 2100 Pennsylvania Avenue, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/416,581B
: FILING DATE: 04-APR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-65825/1994
: FILING DATE: 04-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Nakamura, Dean H.
: REGISTRATION NUMBER: 33,981
: REFERENCE/DOCKET NUMBER: O-37891
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)293-7060

```


TELEFAX: (202)293-7860
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-416-581B-9

Query Match 99.4%; Score 4006; DB 1; Length 770;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 765; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

Oy 1 MAQMNQLOQDLDRYKOLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
Db 1 MAQMNQLOQDLDRYKOLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
Oy 61 LGEIDQOYSRFLQESNVLYOHNLRRIKOFLQSRYLEKPMELIARIYARCLMESSRLQTA 120
Db 61 LGEIDQOYSRFLQESNVLYOHNLRRIKOFLQSRYLEKPMELIARIYARCLMESSRLQTA 120
Oy 121 TAAQGGGANHPTAAVTEKQOMLEQHLQDVRRKVODLEQKKVVENLODDDFENYKTLK 180
Db 121 TAAQGGGANHPTAAVTEKQOMLEQHLQDVRRKVODLEQKKVVENLODDDFENYKTLK 180
Oy 181 SOGDMDLNGNNOYSTROKMOOLEQMLTALDQMRISYSELGLLSAMEYVOKTLTDEEL 240
Db 181 SOGDMDLNGNNOYSTROKMOOLEQMLTALDQMRISYSELGLLSAMEYVOKTLTDEEL 240
Oy 241 ADMKRRPEIACIGPPNICLDRLENNWITSLAESQLOTROQIKKLEELQOKVSKGDPVQ 300
Db 241 ADMKRRPEIACIGPPNICLDRLENNWITSLAESQLOTROQIKKLEELQOKVSKGDPVQ 300
Oy 301 HRPMLEERIVELFRNLMSAFVVEROPCMHPMDRPLVKTGVQFTTKVRLVKKPELVY 360
Db 301 HRPMLEERIVELFRNLMSAFVVEROPCMHPMDRPLVKTGVQFTTKVRLVKKPELVY 360
Oy 361 QLKIKVCIDKDSGDVAALGSRKFNILGNTKVMNNEESNNGSLSAEFKHLTLREORGN 420
Db 361 QLKIKVCIDKDSGDVAALGSRKFNILGNTKVMNNEESNNGSLSAEFKHLTLREORGN 420
Oy 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANMASILMY 480
Db 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANMASILMY 480
Oy 481 NMLTNPKVNFETPRPIGTWDOVAEVLISWOPSSSTTKRGLSTIEOLTTLAEXLLGPGVNS 540
Db 481 NMLTNPKVNFETPRPIGTWDOVAEVLISWOPSSSTTKRGLSTIEOLTTLAEXLLGPGVNS 540
Oy 541 GCOITMAFKCKENMAKGFSFWWMLDNIIDLVKYTLALAMNGYIMGFTSKRERAILST 600
Db 541 GCOITMAFKCKENMAKGFSFWWMLDNIIDLVKYTLALAMNGYIMGFTSKRERAILST 600
Oy 601 KRPGFLLRFSSSSKEGCVTFTWKEKIDSGKTQIOSVEYTTKOOLNMSPFAELIMGYKIM 660
Db 601 KRPGFLLRFSSSSKEGCVTFTWKEKIDSGKTQIOSVEYTTKOOLNMSPFAELIMGYKIM 660
Oy 661 DATNLIIVSLVLYPDIPKEEAFKCYKRPESOBHPADGSAAPYKTFICVPTPTCSN 720
Db 661 DATNLIIVSLVLYPDIPKEEAFKCYKRPESOBHPADGSAAPYKTFICVPTPTCSN 720
Oy 721 TIDLPMSPRTLDLSLMQFGNNGGAEPBSAGQFESLTFMDLTLSECATSPM 770
Db 721 TIDLPMSPRTLDLSLMQFGNNGGAEPBSAGQFESLTFMDLTLSECATSPM 770

```

RESULT 12
 US-08-416-581B-1
 Sequence 1, Application US/08416581B
 Patent No. 5719042
 GENERAL INFORMATION:
 APPLICANT: kishimoto, Tadimitsu

APPLICANT: Akira, Shizuo
 TITLE OF INVENTION: TRANSCRIPTION FACTOR APRF
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
 STREET: 2100 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/416,581B
 FILING DATE: 04-APR-1995
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: JP 6-65825/1994
 FILING DATE: 04-APR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Nakamura, Dean H.
 REGISTRATION NUMBER: 33,981
 REFERENCE/DOCKET NUMBER: 0-37891
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)293-7060
 TELEFAX: (202)293-7860
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 STRANDS: 1
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-416-581B-1

Query Match 98.7%; Score 3977; DB 1; Length 770;
 Best Local Similarity 98.6%; Pred. No. 0;
 Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

Oy 1 MAQMNQLOQDLDRYKOLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
Db 1 MAQMNQLOQDLDRYKOLHQLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNL 60
Oy 61 LGEIDQOYSRFLQESNVLYOHNLRRIKOFLQSRYLEKPMELIARIYARCLMESSRLQTA 120
Db 61 LGEIDQOYSRFLQESNVLYOHNLRRIKOFLQSRYLEKPMELIARIYARCLMESSRLQTA 120
Oy 121 TAAQGGGANHPTAAVTEKQOMLEQHLQDVRRKVODLEQKKVVENLODDDFENYKTLK 180
Db 121 TAAQGGGANHPTAAVTEKQOMLEQHLQDVRRKVODLEQKKVVENLODDDFENYKTLK 180
Oy 181 SOGDMDLNGNNOYSTROKMOOLEQMLTALDQMRISYSELGLLSAMEYVOKTLTDEEL 240
Db 181 SOGDMDLNGNNOYSTROKMOOLEQMLTALDQMRISYSELGLLSAMEYVOKTLTDEEL 240
Oy 241 ADMKRRPEIACIGPPNICLDRLENNWITSLAESQLOTROQIKKLEELQOKVSKGDPVQ 300
Db 241 ADMKRRPEIACIGPPNICLDRLENNWITSLAESQLOTROQIKKLEELQOKVSKGDPVQ 300
Oy 301 HRPMLEERIVELFRNLMSAFVVEROPCMHPMDRPLVKTGVQFTTKVRLVKKPELVY 360
Db 301 HRPMLEERIVELFRNLMSAFVVEROPCMHPMDRPLVKTGVQFTTKVRLVKKPELVY 360
Oy 361 QLKIKVCIDKDSGDVAALGSRKFNILGNTKVMNNEESNNGSLSAEFKHLTLREORGN 420
Db 361 QLKIKVCIDKDSGDVAALGSRKFNILGNTKVMNNEESNNGSLSAEFKHLTLREORGN 420
Oy 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANMASILMY 480
Db 421 GGRANDASLIYTEELHLITFETEVYHOGKIDLETHSLPVVVISNICOMPANMASILMY 480

```

Db 421 GGRANDASLIYTEELHLTFETEVYHOGKIDLETHSLSVVVISNICOMPANASILMY 480
Oy 481 NMLTNPKXVNFPTKPPICITWDOVAEVLISWOFSTTKRGLSIEOLTTTAEKLLGPGVNS 540
Db 481 NMLTNPKXVNFPTKPPICITWDOVAEVLISWOFSTTKRGLSIEOLTTTAEKLLGPGVNS 540
Oy 541 GCOITWAKCKEMAGKGFSEFWWLDNIIDLVKYTLALANEGYIMGFISKEKERAIIIST 600
Db 541 GCOITWAKCKEMAGKGFSEFWWLDNIIDLVKYTLALANEGYIMGFISKEKERAIIIST 600
Oy 601 KPPGTFLLRFSESSKEGVTFTWVKEDISGKTIOISVEPYTKOOLNNMSFAEIIIMGYKIM 660
Db 601 KPPGTFLLRFSESSKEGVTFTWVKEDISGKTIOISVEPYTKOOLNNMSFAEIIIMGYKIM 660
Oy 661 DATNIIIVSLVLYLPDIPEEAFGKICRPSOEHPRADGSAAPYLKTFICVPTTCSN 720
Db 661 DATNIIIVSLVLYLPDIPEEAFGKICRPSOEHPRADGSAAPYLKTFICVPTTCSN 720
Oy 721 TIDLPMSPTLDSLMQFGNNGGAEPSAGOFESLTFDMDLTSECATSPM 770
Db 721 TIDLPMSPTLDSLMQFGNNGGAEPSAGOFESLTFDMDLTSECATSPM 770

RESULT 13
US-08-416-581B-5
: Sequence 5, Application US/08416581B
: Patent No. 5719042
: GENERAL INFORMATION:
: APPLICANT: Kishimoto, Tadamitsu
: APPLICANT: Akita, Shizuo
: TITLE OF INVENTION: TRANSCRIPTION FACTOR APF
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
: STREET: 2100 Pennsylvania Avenue, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/416,581B
: FILING DATE: 04-APR-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-65825/1994
: FILING DATE: 04-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Nakamura, Dean H.
: REGISTRATION NUMBER: 33,981
: REFERENCE/DOCKET NUMBER: O-37891
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)293-7060
: TELEFAX: (202)293-7860
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 770 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-416-581B-5

Query Match 98.7%; Score 3977; DB 1; Length 770;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
Oy 1 MAOMNLOOQLODTRVYLKQLHOLVSDTFPMELRQFLAPWIESQDMAYAAKESHATLVFHNL 60
|||||

Db 1 MAOMNLOOQLODTRVYLKQLHOLVSDTFPMELRQFLAPWIESQDMAYAAKESHATLVFHNL 60
Oy 61 LCEIDOOVSRLFOESNVLYOHNLRLRIKOFLOSRYLEKPMELARIYARCLMESSRLLOTA 120
Db 61 LCEIDOOVSRLFOESNVLYOHNLRLRIKOFLOSRYLEKPMELARIYARCLMESSRLLOTA 120
Oy 121 TAAOOGGQANHPPTAAVVFEEKOQMLEOHLODVKR RVODI.EOKKVVENLODDDFPNVKTIL 180
Db 121 TAAOOGGQANHPPTAAVVFEEKOQMLEOHLODVKR RVODI.EOKKVVENLODDDFPNVKTIL 180
Oy 181 SOGDMODLGNNOVSYTRKMOOLEQMLTALDOMRSIVSELAGLLSAMEYVOKTILTDEEL 240
Db 181 SOGDMODLGNNOVSYTRKMOOLEQMLTALDOMRSIVSELAGLLSAMEYVOKTILTDEEL 240
Oy 241 ADMKRRPQLACIGGPNICIDRLNMTSLASOLOTRQOIKKLELOOKVSKDPITVQ 300
Db 241 ADMKRRPQLACIGGPNICIDRLNMTSLASOLOTRQOIKKLELOOKVSKDPITVQ 300
Oy 301 HRPMLEERIVLELFRNLMSAFYVEROPCPMPHDPDRP.VIKTGVQFTTKVRL.VKPEELNY 360
Db 301 HRPMLEERIVLELFRNLMSAFYVEROPCPMPHDPDRP.VIKTGVQFTTKVRL.VKPEELNY 360
Oy 361 QLKIKVCIDKSDGVAALGSKRFNLTGNTKYMMNEESNNSLSAEKHLTLREORGN 420
Db 361 QLKIKVCIDKSDGVAALGSKRFNLTGNTKYMMNEESNNSLSAEKHLTLREORGN 420
Oy 421 GGRANDASLIYTEELHLTFETEVYHOGKIDLETHSLPVVVISNICOMPANASILMY 480
Db 421 GGRANDASLIYTEELHLTFETEVYHOGKIDLETHSLPVVVISNICOMPANASILMY 480
Oy 481 NMLTNPKXVNFPTKPPICITWDOVAEVLISWOFSTTKRGLSIEOLTTTAEKLLGPGVNS 540
Db 481 NMLTNPKXVNFPTKPPICITWDOVAEVLISWOFSTTKRGLSIEOLTTTAEKLLGPGVNS 540
Oy 541 GCOITWAKCKEMAGKGFSEFWWLDNIIDLVKYTLALANEGYIMGFISKEKERAIIIST 600
Db 541 GCOITWAKCKEMAGKGFSEFWWLDNIIDLVKYTLALANEGYIMGFISKEKERAIIIST 600
Oy 601 KPPGTFLLRFSESSKEGVTFTWVKEDISGKTIOISVEPYTKOOLNNMSFAEIIIMGYKIM 660
Db 601 KPPGTFLLRFSESSKEGVTFTWVKEDISGKTIOISVEPYTKOOLNNMSFAEIIIMGYKIM 660
Oy 661 DATNIIIVSLVLYLPDIPEEAFGKICRPSOEHPRADGSAAPYLKTFICVPTTCSN 720
Db 661 DATNIIIVSLVLYLPDIPEEAFGKICRPSOEHPRADGSAAPYLKTFICVPTTCSN 720
Oy 721 TIDLPMSPTLDSLMQFGNNGGAEPSAGOFESLTFDMDLTSECATSPM 770
Db 721 TIDLPMSPTLDSLMQFGNNGGAEPSAGOFESLTFDMDLTSECATSPM 770

RESULT 14
US-09-087-465-6
: Sequence 6, Application US/09087465A
: Patent No. 6164992
: GENERAL INFORMATION:
: APPLICANT: Vinkemeier, Uwe
: APPLICANT: Chen, Xiaomin
: APPLICANT: Darnell Jr., James E
: APPLICANT: Kurtyan, John
: TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS OF
: FILE REFERENCE: 600-1-229
: CURRENT APPLICATION NUMBER: US/09/087,465A
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 6
: LENGTH: 770
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-087-465-6

```

Query Match          98.7%; Score 3977; DB 4; Length 770;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAOMNLOOQDTRYKOLHOLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNH 60
DB 1 MAOMNLOOQDTRYKOLHOLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNH 60
OY 1 LGELIDQYSFLEDSNVLYOHNLRIKOPLOSRYLEKPMELARIVARCLMEESRLLOTAA 120
DB 1 LGELIDQYSFLEDSNVLYOHNLRIKOPLOSRYLEKPMELARIVARCLMEESRLLOTAA 120
OY 121 TAAOOGGQANHPTAAVYTEKQOMLEOHLQDVRKRVODLEOKMKVENVLQDDPFNYKTKL 180
DB 121 TAAOOGGQANHPTAAVYTEKQOMLEOHLQDVRKRVODLEOKMKVENVLQDDPFNYKTKL 180
OY 181 SOGDMODLNNNSVTRKQKQOLEOMLTALDOMRSTIVSELSLAMEYVOKTLTDEEL 240
DB 181 SOGDMODLNNNSVTRKQKQOLEOMLTALDOMRSTIVSELSLAMEYVOKTLTDEEL 240
OY 241 ADMKRRPEIACIGPPNICDLRLNNITSLAESOLQTRQOIKKLEELQOKVSKGDPYVQ 300
DB 241 ADMKRRPEIACIGPPNICDLRLNNITSLAESOLQTRQOIKKLEELQOKVSKGDPYVQ 300
OY 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVITGVQFTTKVRLVKFPELNY 360
DB 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVITGVQFTTKVRLVKFPELNY 360
OY 361 QLKIKVCIDKSGDVAALRSGRKNILGTNTKYNNMEESNNGSLSAEFKHLTREORCGN 420
DB 361 QLKIKVCIDKSGDVAALRSGRKNILGTNTKYNNMEESNNGSLSAEFKHLTREORCGN 420
OY 421 GGRANDASLIYTEELHLLTFETEVYHOGKIDLETHSLPLVVIISNICOMPANASILMY 480
DB 421 GGRANDASLIYTEELHLLTFETEVYHOGKIDLETHSLPLVVIISNICOMPANASILMY 480
OY 481 NMLTNNPKVNFPTKPRIGTMDOVAEVLWSQFSSTTKRGISTEOLTTLAKKLGPGVNS 540
DB 481 NMLTNNPKVNFPTKPRIGTMDOVAEVLWSQFSSTTKRGISTEOLTTLAKKLGPGVNS 540
OY 541 GCOITWAKFCCKENMAGKGSFWWLDNIIDLVKYILALMNEGYIMGFISKEBERA1LST 600
DB 541 GCOITWAKFCCKENMAGKGSFWWLDNIIDLVKYILALMNEGYIMGFISKEBERA1LST 600
OY 601 KPPGTFLLRFSSSKEGVTFTEWEDISGKTQIOSVEPTTKOOLNMSFAELIMGKIM 660
DB 601 KPPGTFLLRFSSSKEGVTFTEWEDISGKTQIOSVEPTTKOOLNMSFAELIMGKIM 660
OY 661 DATNILVSPLYLVPDIKPEEAFGKYCRPESQEHPEADPGSAAPYLKTFICVTPPTCSN 720
DB 661 DATNILVSPLYLVPDIKPEEAFGKYCRPESQEHPEADPGSAAPYLKTFICVTPPTCSN 720
OY 721 TIDLPMSPRALDLSLMQFGNNGEGAEPSAGQFESLTFDMELTSECATSPM 770
DB 721 TIDLPMSPRALDLSLMQFGNNGEGAEPSAGQFESLTFDMELTSECATSPM 770

RESULT 15
US-08-276-099A-14
Sequence 14, Application US/08276099A
Patent No. 5591825
GENERAL INFORMATION:
APPLICANT: McKnight, Steven L
APPLICANT: Hou, Jinzhao
TITLE OF INVENTION: INTERLEUKIN-4 SIGNAL TRANSDUCERS AND
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSER: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California

```

```

; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276, 099A
; FILING DATE: 15-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59451-1/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELETYPE: 910 277299
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-276-099A-14

```

```

Query Match          98.7%; Score 3977; DB 1; Length 771;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

```

OY 1 MAOMNLOOQDTRYKOLHOLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNH 60
DB 1 MAOMNLOOQDTRYKOLHOLYSDTFPMELROFLAPWIESODMAVAASKESHATLVFHNH 60
OY 1 LGELIDQYSFLEDSNVLYOHNLRIKOPLOSRYLEKPMELARIVARCLMEESRLLOTAA 120
DB 1 LGELIDQYSFLEDSNVLYOHNLRIKOPLOSRYLEKPMELARIVARCLMEESRLLOTAA 120
OY 121 TAAOOGGQANHPTAAVYTEKQOMLEOHLQDVRKRVODLEOKMKVENVLQDDPFNYKTKL 180
DB 121 TAAOOGGQANHPTAAVYTEKQOMLEOHLQDVRKRVODLEOKMKVENVLQDDPFNYKTKL 180
OY 181 SOGDMODLNNNSVTRKQKQOLEOMLTALDOMRSTIVSELSLAMEYVOKTLTDEEL 240
DB 181 SOGDMODLNNNSVTRKQKQOLEOMLTALDOMRSTIVSELSLAMEYVOKTLTDEEL 240
OY 241 ADMKRRPEIACIGPPNICDLRLNNITSLAESOLQTRQOIKKLEELQOKVSKGDPYVQ 300
DB 241 ADMKRRPEIACIGPPNICDLRLNNITSLAESOLQTRQOIKKLEELQOKVSKGDPYVQ 300
OY 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVITGVQFTTKVRLVKFPELNY 360
DB 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPDRPLVITGVQFTTKVRLVKFPELNY 360
OY 361 QLKIKVCIDKSGDVAALRSGRKNILGTNTKYNNMEESNNGSLSAEFKHLTREORCGN 420
DB 361 QLKIKVCIDKSGDVAALRSGRKNILGTNTKYNNMEESNNGSLSAEFKHLTREORCGN 420
OY 421 GGRANDASLIYTEELHLLTFETEVYHOGKIDLETHSLPLVVIISNICOMPANASILMY 480
DB 421 GGRANDASLIYTEELHLLTFETEVYHOGKIDLETHSLPLVVIISNICOMPANASILMY 480
OY 481 NMLTNNPKVNFPTKPRIGTMDOVAEVLWSQFSSTTKRGISTEOLTTLAKKLGPGVNS 540
DB 481 NMLTNNPKVNFPTKPRIGTMDOVAEVLWSQFSSTTKRGISTEOLTTLAKKLGPGVNS 540
OY 541 GCOITWAKFCCKENMAGKGSFWWLDNIIDLVKYILALMNEGYIMGFISKEBERA1LST 600
DB 541 GCOITWAKFCCKENMAGKGSFWWLDNIIDLVKYILALMNEGYIMGFISKEBERA1LST 600

```


xx Receptor recognition factor implicated in transcriptional
 PT stimulation of genes - useful in drug screening assays and/or
 PT for treating cellular debilities, derangements and/or
 PT dysfunctions, etc.
 PS Claim 1; Page 107-110; 160pp; English.
 xx A fragment encoding the human Stat91 protein was used to screen a
 CC murine thymus and spleen cDNA for homologous proteins. A highly
 CC homologous gene (given in AA089338) was isolated that encoded a
 CC 91 kDa protein (AA072080) (Stat1) that was responsive to interferon-
 CC gamma. Using a fragment of the mouse gene as probe, 2 additional
 CC members of the 113-91 family of receptor recognition factor
 CC proteins were isolated. The 2 genes (AA089339-40) were cloned
 CC in plasmids 13sf1 and 19sf6 and encoded proteins termed Stat4
 CC (AA072081) and Stat3 (AA072082), respectively.
 xx
 SQ Sequence 770 AA;

Query Match 100.0%; Score 4029; DB 16; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAOMNLOQLDTRYLKOLHLYSDTFPMELROFLAPWIESQPMAYAAKESHATLVFHNL 60
 DB 1 maqmqnlgldtrylkqhlqlysdtfpmlrqlapwiesqgwajaaakeshatlvlfnl 60
 QY 61 LGEIDQVSRFLQESNVLYOHNLRRIKOFLQSRYLEKPEIARIYARCLWESRLLOTAA 120
 DB 61 lgeidqysrflqesnvlyqhnllrrikqflqsrylekpeiarivarclewesrlllqtaa 120
 QY 121 TAAOOGGANHPHTAVNTEKQOMLEOHLODYRKRVODLEQKKVYENLODDDFNFKTK 180
 DB 121 taaogqganhphtaavnteekqomleqhldyrvkrvqdlqekkvvevnloddidfnkklk 180
 QY 121 caaqqgqanhphtaavnteekqomleqhldyrvkrvqdlqekkvvevnloddidfnkklk 180
 DB 121 caaqqgqanhphtaavnteekqomleqhldyrvkrvqdlqekkvvevnloddidfnkklk 180
 QY 181 SOGDMDLNGNNSVTRKQMOLEOMLTALDOMRSIVSELGKLISAMEYOKTLTDEEL 240
 DB 181 sqgdmdlgnngsvtrkqmqleqmtaldomrsivselagllsameyoktltdel 240
 QY 181 sqgdmdlgnngsvtrkqmqleqmtaldomrsivselagllsameyoktltdel 240
 DB 181 sqgdmdlgnngsvtrkqmqleqmtaldomrsivselagllsameyoktltdel 240
 QY 241 ADWKRREPIACIGPPNICLDRLENWITSLSAQSLOTROQIKKLELOQVSYKGPPIVO 300
 DB 241 adwkrrepiaci gppnicldrlenwitslaesqlqtrqikklelqkvsykgpplvq 300
 QY 301 HRPMLREIVLEFRMLKSAFVERQPCMPHDPRLVKTGVOTTKRLLVKKPELNY 360
 DB 301 hrpmlreivlefrmlksafverqpcmpmhdpriplvktgvttkrvllvkkpeln 360
 QY 361 QLKTKVCLDKDSGVAAALGRSKRFNLIOTNTKVMNNEESNNGSLSAEFKHLTLRQRCGN 420
 DB 361 qlktkvcldkdsdvaaalgrsrkfnllgtlncvmneesngslsaekhlhllrqrcgn 420
 QY 421 GGRANCDASLIVTEELHITFETEVYHOGCLKIDLEHSLPVVYVINSICOMPANASILMY 480
 DB 421 ggrancdaslivteelhltfetyvhogclkidlehsllpvvynslcqpnaasilly 480
 QY 481 NMLTNNPNVNFETPPICGTDOVAEVLNWSQSTTKRGLSTEOLTTLAEXLLGGVNS 540
 DB 481 nmltnnpnvnfetppicgtwdaevlswqsttkrglstlqlllaeklllpgvny 540
 QY 541 GCQITWAKFCKENNAKGFSFWWLDNITLVKXYILALMNEGYIMGFSKREERAILST 600
 DB 541 gcqitwakfckennakgfsfwvldnildlvkxyilalmnwyimgfiskerailst 600
 QY 601 KPRCFILRFSESSKEGCVTFMWKDISKQTQIOSVEPYTQOULNNNSFAITMGYKIM 660
 DB 601 kprcfilrfseesskegcvtfmwkdisqtkqisvepytkqulnnnsfaeilmgykim 660
 QY 661 DATNILLSPLYLYPDJPKBEAFGKYCRPESQHPBADGSAAPYLKTKFICVTPTCSN 720
 DB 661 datnillvspllylypdjpkbeafgkycrpesqhpbadgsaapylktkficvptcsn 720
 QY 721 TIDLPMSPRTLDLMOFCNNNGECAPSGOPESTLFFDMDLTSECATSPM 770
 DB 721 tidlpmsprrtldlmo fcnnngecapsgopestlffdmldtsecatspm 770

QY 721 TIDLPMSPRTLDLMOFCNNNGECAPSGOPESTLFFDMDLTSECATSPM 770
 DB 721 tidlpmsprrtldlmo fcnnngecapsgopestlffdmldtsecatspm 770

RESULT 2

AAW03176
 ID AAW03176 standard; Protein; 770 AA.
 AC AAW03176;
 DT 24-OCT-1996 (first entry)
 XX
 DE Mouse STAT4.
 KW STAT; STAT4; signal transducer and activator of transcription;
 KW DNA binding protein; ligand; receptor; oncogenesis; inflammation;
 KW autoimmune disease; antagonist; therapy.
 OS Mus sp.
 FH Key Location/Qualifiers
 FT Domain 398..508
 FT /label= DNA_binding_domain
 FT /note= "Claim 3, page 110"

W09620954-A2.

11-JUL-1996.

28-DEC-1995; 95WO-US17025.

06-JAN-1995; 95US-0369796.

(UYRO) IV ROCKEFELLER.

Darnell JE, Horvath CM, Wen Z, Zhong Z;

WPI: 1996-333941/33.

N-PSDB; AAT31280.

xx New STAT protein DNA-binding domain peptide(s) - useful for
 PT diagnosis, preventing or treating cellular dysfunction, e.g.
 PT oncogenesis, inflammation, parasitic disease or autoimmunity
 PS Disclosure: Page 87-90; 138pp; English.

xx Mouse signal transducer and activator of transcription (STAT)
 CC protein STAT4 (AAW03176) serves a dual purpose, i.e. signal
 CC transduction from ligand-activated receptor kinase complexes
 CC followed by nuclear translocation and DNA binding to activate
 CC transcription. Recombinant STAT4 can be obtd. using cDNA clone
 CC 19sf6 (AAT31278) obtd. from splenic/thymic cells. STAT4 includes
 CC a DNA-binding domain (see also AAW03167) capable of both
 CC receptor recognition and message delivery via DNA binding in a
 CC receptor-ligand specific manner. STAT proteins and their DNA
 CC binding domains (see also AAW03165-75) are useful for screening
 CC antagonists used to inhibit STAT-mediated signal transduction
 CC and activation of transcription.

xx Sequence 770 AA;

Query Match 100.0%; Score 4029; DB 17; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAOMNLOQLDTRYLKOLHLYSDTFPMELROFLAPWIESQPMAYAAKESHATLVFHNL 60
 DB 1 maqmqnlgldtrylkqhlqlysdtfpmlrqlapwiesqgwajaaakeshatlvlfnl 60
 QY 61 LGEIDQVSRFLQESNVLYOHNLRRIKOFLQSRYLEKPEIARIYARCLWESRLLOTAA 120
 DB 61 lgeidqysrflqesnvlyqhnllrrikqflqsrylekpeiarivarclewesrlllqtaa 120

Db	61	lgeIdqystrflqgesnvlqghnllrtlkqflqsfylyiekpmeariavrciweesrlllqta	120
QY	121	TAADGQGANHPHTAAVYTEKQOMLEOHLODVRRKRVODLEQKMKVENVLODFDFNYKTLK	180
Db	121	laaqqgganhrptaavvtekeqmqleghllqdvtrkvrqqlleqkmkvvlenlqddfdafnykllk	180
QY	181	SQGMMDLNCNNSVTRQKKMOOLEOMLTALDOMKRSIVSELAGLSAMEVYQKTLTDEEL	240
Db	181	sqgmqdgdnngnsvtrcqkmgqlleqmltalqdmrslsvselaqljsamevvgkllldee	240
QY	241	ADMKRPEEACIGCPNPNICDLRELNMTTSLAESQLOQTRQOIKKLEELQOKVYKSGDPYQ	300
Db	241	admkrrpeeaciagppnicldrlenwltslaesqlqtrqgikkleelqkvsvykgdpivq	300
QY	301	HRPMLEERTIVLEFNNLKKSAFVVRQOCMPHRRPLVITGVQFTTKVALLKPEELNT	360
Db	301	hrpmleerivlelfnlnksaflvrerqocmhprrplvrltkyqftkvlvlpkfdelnv	360
QY	361	QLKIKVCIDKDSGDVAALRGSRKFNITLGNTMKVYNNMEESNNGSLSAEFKHLTLREORGN	420
Db	361	qlkikvcidksgdvaalrgsrkfnllgntmkymmeesnngslsaefkhlrlregrcn	420
QY	421	GGRANCDASLIVTEBELHILITRETVYHQGLIDLETHSLPYYVVISNCOMPNAMASILMT	480
Db	421	ggrancdasllivteelhlitfetevyhqglikidletlshlpvvyvisncompnawasllwy	480
QY	481	NMLTNPNKVNVEFFKPRPGTMDQVAEVLDSMOFSSTTKRGSLIEDLTLAEKLLGPGVNY	540
Db	481	nmltnpknvfflcpkprlgtdqvaevlswqfssctlrgrjsieqltlleakllpgpyvny	540
QY	541	GCOITWAKFCKENAGKGFSEFWMLDNTIDLKVKYILALNNEGYIMGFISKEREALIST	600
Db	541	gcgitwakfckennaqgfsfwwldnldlvtkkylalawmegyimgfiskereralist	600
QY	601	KPPGTFLLARESESSKEGVTFTTWEXKOISGKTQIQSEVPTTKQOLNNKMSFAELIIMGKIM	660
Db	601	kppgtflltrisessekegvtfttwexkolsqtklqsvaypkykqqlnmstfaelimgykm	660
QY	661	DATNITLVSPLYLVLPDIPKEAEFAKCYCRPESOEHPREADPGSAAPYLTKFTICVPTTCGN	720
Db	661	datnllvsplvlylpdpkreaefakycrcpseohpreadpgsaapylltkfticvpttcn	720
QY	721	TIDLPMSRRTIDSLMOGNCNEGAEPSAGGOFESLFTDMDLTSCATSPM	770
Db	721	tldlpmsprtlidslmqfngnqgaepsagqfieslftdmiltsecatspm	770
RESULT 3			
AABI2377			
ID	AABI2377 standard; peptide: 770 AA.		
XX			
AC	AABI2377:		
XX			
DT	08-NOV-2000 (first entry)		
XX			
DE	N-terminal domain of murine STAT-3 protein.		
XX			
KW	STAT; signal transducer and activator of transcription; crystal;		
RW	drug design; murine.		
OS	Mus sp.		
XX			
FH	Key	Location/Qualifiers	
FT	Region	4..9	/label= "Alpha helix 1"
FT	Region	12..21	/label= "Alpha helix 2"
FT	Region	19..21	/label= "Alpha helix 2"
FT	Region	28..33	/label= "3(10) helix of alpha helix 2"
FT	Region	35..40	/label= "Alpha helix 3"
FT	Region	35..40	/label= "Alpha helix 4"

Query	Match	Best Local Similarity	Score	DB	Length	770:
Matches	768:	Conservative	1:	Mismatches	1:	Indels 0: Gaps 0:
QY	1	MAOMNLOOLDIRRYKOLHOLSDIFPEMLROFLAPMIESQOMAWAASKESHATLVFHNL	60			
DB	1	mqgwmgqlgldrrykgqlhqlglsdfpmeirgflapwlesqgwayaaskeshatlvfnnl	60			
QY	61	IGELIDQVSRFLQESNVLYOHNLRRIRIKOFLOSRYLEKPMELIARIYAKLWESRLQTTAA	120			
DB	61	lgeidqlysrflqesnvllyghnlrrirkfllgrylekpmelariyarcliweesrltqtaa	120			
QY	121	TAAOCCGQANHPATAVVRKEKQOMLEOHLODYVKRRQDLEOKKKVVENLODDPDMYKTLK	180			
DB	121	taagqggaanhpataavaavlekqgmleqhnqdvkrtyqdlqgmkvvenlqddtdinykclk	180			
QY	181	SOGDMODLGNNOQSYTRKMOQLEOMLTALDQMRKSIVSELAGLISAMEYVOKTLTDEEL	240			
DB	181	sggdmgdlnngnqsvtrtqmgqlqemltaldqmrsviseagllsameyqkltideel	240			
QY	241	ADMKRRPEELACIGGPPNICTLDRLNEMWITISLASQLOTRQIKKLEEDQKSYKQDPYVO	300			
DB	241	admkrrpaelaciggppnictldrlnemwitislaesqilqtrqikkleeiqkvsykgdpyiv	300			
QY	301	HHPMLEERIVLEFRNLKMSAFVVERQPCMPHDPBPVLKGTGQPTTKVRLLVKRPPELNY	360			
DB	301	hnpmlleerivlefrnlkmsafvverqpcmpmhdpvlykgtvgfttkvrrllvkrppei	360			
QY	361	QKIKVIVIDKDSGDVAALGSRKRNFIILCTNTRKVNMEESNNGSLSAEERKHLTLREORCGN	420			

Db 361 qlkikvcldksdgvaaalrgs-rkfnlglntkvmmeesnglsjaefkhlilreqrcgn 420
QY 421 GGRANDASLIYTFEELHILTFEVEYHOGIAKIDLETHSLPVAVINSICOMPANASILMY 480
Db 421 ggrandasllyteelhlitfetevyhgkikidletshlpvvnslcgmpanasllwy 480
QY 481 NMLTNPKNVNFYTKPRIGTMDVAEVLISWOFSSSTTKRGLSIEQLTTLAEKLLGCVNYS 540
Db 481 nmltnpkvnftfkprlgtwqvaevlswqfsssttkrglsieqlttlaekllpgvny 540
QY 541 GGOITWAKFCENMAGKGSFWWMLDNIIDLVKKYILALMNGYIMGFISKREERAILST 600
Db 541 ggoitwakfckenmagkgsfwvwlidnlvkkylalwngyimgfiskerailstc 600
QY 601 KPPGTFLLRFSSSKEGVTFWVEKDISGKTQIQSVEPYTRKQOLNNMSFAEIMGYKIM 660
Db 601 kppgtfllrfssskegytlfwvekdlsqkctqigsvepytkqqlnmsfaelimgyim 660
QY 661 DATNIIIVSPVLYLPDIPKEEAFGKCYCRPESOEHPADGSAAPYIKTFICVPTTCN 720
Db 661 datnllivsplylylpdipkeeafigkycrpesqehpeadpysaapyiktkfctvptc 720
QY 721 TIDLPMSPRTLDLMOFGNNGCAEPSAGQFESLTFDMDLTSECATSPM 770
Db 721 tidlpmsprtldslmqfngngeaepsagqfesltfdmdltsecatspm 770

RESULT 4
AAR82995
ID AAR82995 standard: Protein: 770 AA.
AC AAR82995:
XX 25-MAR-1996 (first entry)
XX Mouse liver acute phase response factor.
DE Mouse liver acute phase response factor.
XX
XX Mouse: acute phase response factor: transcription factor;
KW Interleukin-6; signal transmission; liver; antibody; antisense;
KW ribozyme; antinflammatory; antitumor; hypotensive; therapy.
XX
OS Mus musculus.
XX
PN EP676469-A2.
XX
PD 11-OCT-1995.
XX
PF 29-MAR-1995; 95EP-0104670.
XX
PR 04-APR-1994; 94JP-0065825.
XX
PA (KISH/) KISHIMOTO T.
XX
PI Akira S, Kishimoto T;
XX
DR WPI: 1995-346089/45.
DR N-PSDB: AAT05619.
XX
PT New acute phase response factor - for developing inhibitory agents
PT for treating diseases induced by cytokine(s) such as IL-6, e.g.
PT inflammatory diseases
XX
PS Claim 10; Page 20-22; 31pp: English.
XX
XX The sequence represents a mouse acute phase response factor (APRF),
XX a transcription factor related to signal transmission of
XX interleukin-6 (IL-6). The protein is encoded by a cDNA, isolated
XX from a mouse liver cDNA library using a polymerase chain reaction
XX product (amplified using primers derived from an IL-6-treated mouse
XX liver peptide) as a probe. APRF-inhibitors, e.g. antibodies,
XX antisense oligonucleotides or ribozymes, may be used to treat
XX diseases induced by IL-6, e.g. inflammatory disease, leukemia,
XX cancer, osteoclastia, pulmonary hypertension, etc.

XX SQ Sequence 770 AA:
Query Match 99.4%; Score 4006; DB 16; Length 770;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 765; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAOWNQLOLDTRYLKQHLQLYSDTFPMELRQFLAPWIESQDMAYAAKESHATLVFHNL 60
Db 1 maqwnqlqldtryleqlhqlysdsfpmelrqflapwiesqdwaaakeshatlvfhn 60
QY 61 LGEIDQYSRELIQESNVLYQHNLRRIKOFLOSRYLEKPMELIARIYARCLMESSRLLOTA 120
Db 61 lgeidqysrliqesnvlyqhnllrrikqflosrylekpmelariyarcclmewsrllqta 120
QY 121 TAAOOGGQGNHPTAAVAVTEKQOMLEOHLODVKRRVODLEQKKKVVENLODDDFNYKTLK 180
Db 121 taagggggnhptaavavtekgqmlqehldvkrkrvqdeqkkkvvnenloddfrnykllk 180
QY 181 SOGDMODLGNNGSVTRKQMQLQOMLTAQDMRNSIVSELAGLSAMEYVOKTLTDEEL 240
Db 181 sogdmmdlgnngsvtrkqmqleqmltalqdmrnsivselagllsamyvqkltldee 240
QY 241 ADMKRRPEIACIGSPNICLDRLENNITSLAESQLOTROQIKLELQOKVSKGDPVYQ 300
Db 241 admkrrpqiacigspnicldrlennitslaesqlotrqoikllelqkvskgdpvlyq 300
QY 301 HRPMLEERIVLELFRNLKSAFVVERQPCMPMHPDRPLVTKGVQFTTKVRLVKPEELNY 360
Db 301 hrpml eerivlelfrnlksafvverqpcmpmhpdrplvltkvgvfttkvrlvkpeelny 360
QY 361 QLKIKYCIDKSDGVAALALGSRKFNILGNTKVMNNEESNNGSLSAEKNHLLRQRCGN 420
Db 361 qlkikvcldksdgvaaalrgs-rkfnlglntkvmmeesnglsjaefkhlilreqrcgn 420
QY 421 GGRANDASLIYTFEELHILTFEVEYHOGIAKIDLETHSLPVAVINSICOMPANASILMY 480
Db 421 ggrandasllyteelhlitfetevyhgkikidletshlpvvnslcgmpanasllwy 480
QY 481 NMLTNPKNVNFYTKPRIGTMDVAEVLISWOFSSSTTKRGLSIEQLTTLAEKLLGCVNYS 540
Db 481 nmltnpkvnftfkprlgtwqvaevlswqfsssttkrglsieqlttlaekllpgvny 540
QY 541 GGOITWAKFCENMAGKGSFWWMLDNIIDLVKKYILALMNGYIMGFISKREERAILST 600
Db 541 ggoitwakfckenmagkgsfwvwlidnlvkkylalwngyimgfiskerailstc 600
QY 601 KPPGTFLLRFSSSKEGVTFWVEKDISGKTQIQSVEPYTRKQOLNNMSFAEIMGYKIM 660
Db 601 kppgtfllrfssskegytlfwvekdlsqkctqigsvepytkqqlnmsfaelimgyim 660
QY 661 DATNIIIVSPVLYLPDIPKEEAFGKCYCRPESOEHPADGSAAPYIKTFICVPTTCN 720
Db 661 datnllivsplylylpdipkeeafigkycrpesqehpeadpysaapyiktkfctvptc 720
QY 721 TIDLPMSPRTLDLMOFGNNGCAEPSAGQFESLTFDMDLTSECATSPM 770
Db 721 tidlpmsprtldslmqfngngeaepsagqfesltfdmdltsecatspm 770
RESULT 5
AAV03768
ID AAV03768 standard: Protein: 770 AA.
AC AAV03768:
XX
XX 11-JUN-1999 (first entry)
XX Human STAT3 allelic variant.
DE
XX Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;
KW Intracellular transcription factor; Interleukin-6; medicament; variant;
KW

KW pharmaceutical; autoimmune disease; inflammatory; human.
 XX
 OS Homo sapiens.
 XX
 PN EP905234-A2.
 XX
 PD 31-MAR-1999.
 XX
 PF 18-FEB-1998; 98EP-0102774.
 XX
 PR 16-SEP-1997; 97EP-0116061.
 XX
 PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
 PI Della Pietra L, Serlupi-Crescenzi O;
 DR MPI: 1999-192664/17.
 DR N-PSDB; AAX29281.
 XX
 PT New human signal transducer and activator of transcription 3 (STAT3)
 PT allelic variant useful for treatment of autoimmune and inflammatory
 PT disease
 XX
 PS Claim 2; Page 9-13; 32pp; English.
 XX
 CC The present sequence represents a predominant allelic variant of human
 CC signal transducer and activator of transcription 3 (STAT3) protein, an
 CC intracellular transcription factor which mediates IL-6 signals. The
 CC encoding sequence differs from the original published human STAT3 gene
 CC sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3
 CC DNA molecule can be used for the recombinant expression of the variant.
 CC STAT3 protein is useful as a medicament or pharmaceutical composition for
 CC treatment of autoimmune or inflammatory diseases.
 CC
 SQ Sequence 770 AA:
 Query Match 99.4%; Score 4006; DB 20; Length 770;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 765; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MAQNOLOQDTRYLKOLHLYSDTPMELROFLAPIESODMAYAAKSHATLVFHNL 60
 DB 1 maqnmqgqddtlylqhlqjysdfpmeirqlflapiesqdwaaaskshalvhn1 60
 OY 61 LGEIDQOYSFLEDSNVLYQHNLRLIKQFLQSRYLEKPMELIARIVARCLMEESRLQTA 120
 DB 61 lgeidqysrflqesnvlyqhnrlrikqlfqsrylekpmelariarclweesrlqtaa 120
 OY 121 TAAOOGGQANHPRAAVVTEKQMLEQHLQDVKKRVQDLEQKMKVVENLQDDPFPNRYTLK 180
 DB 121 taaqgqgqanhrpaaavvteqqlqhlqdvkrxvqdlqdkmkvvenlqddfdfnkylk 180
 OY 181 SOGDMDLNGNNSVTRKQKQLEQMLTALDQKRSTIVSLAGLSAMEYVOKLTDEEL 240
 DB 181 sqgdmqdlngnsvtrkqmqlegmltalqdmrtisvslaglsameyvkltdeel 240
 OY 241 ADMKRPETACIGCPNLCIDLENNITSLAESQLQTRQOIKKLEELQOVXSKGDPYV 300
 DB 241 admkrrpqiaicgppnicidrlennitslaesqlqtrqgikkleelqovxskgdpivg 300
 OY 301 HRPMLEERIVELFRNLKMSFVVERPCPMHRDRPLVITGVQFTTKVNLVYKPEPLNY 360
 DB 301 hrpml eerivel frnlmksa fverpcpmp hrdrplv itgvqfttkvnlv kfpelny 360
 OY 361 QLKIKVICDKSGDVAALGSRKFNLIGNTKVMNEESNGLSSEFKLTLREOQCN 420
 DB 361 qlkikvicdksgdvaalgsrkfntligntkvmneesnglssefkhltlreqcgn 420
 OY 421 GGRANDASLVTTELHLTFETEVYHQGLKIDLETHSLPVVVISINICQPNMASTILMY 480
 DB 421 ggrandaslvttehlhtfetevyhgkldlethslpvvvisinicqpnmaslilwy 480

OY 481 NMLTNNPKNVNEFTKPPIGTWQVAEVLWSQFSSTTKRGSLIEQLTTLAEKLGPGVNS 540
 DB 481 nmltnnpknnvneftkppigtwqvaevlwsqfsttkrgslieqltllaeklpgpvnys 540
 OY 541 GCQITWAKFCKENMAGKGFSPVWMLDNIIDLVKKTIIALMNEGYIMGTISKREKALLST 600
 DB 541 gcqitwakfckemagkgsfvpwldnidlvkktii al mneyim gtsk rera llst 600
 OY 601 KPPGTELLRFSESSKEGAVTFWTEKDISGKTQIOQSVPEPYTKQOOLNNNSPFEIIMGYKIM 660
 DB 601 kppgtellrfseesskegavtfwtekd is gktqio qsvpepytkqo ol nns pfe iim gykim 660
 OY 661 DATNLSPIVLYLPDIPKEAFKRCRPEQEHPEADPGSAAPYLKTKFTICVPTPTCSN 720
 DB 661 datnlsvpilylpdipkeefgkrcrpesqehpeadpgsaapylkktkfticvptptcsn 720
 OY 721 TIDLPMSPTLDSLMQGNNGEAPPSAGGFESLTFPMDLTSECATSPM 770
 DB 721 tidlpmsptl dslmqgnngeapsaggfesl tfpmdltsecat spm 770
 RESULT 6
 AAB58442
 ID AAB58442 standard; Protein: 793 AA.
 XX
 AC AAB58442;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 780.
 XX
 KW Human; lung cancer associated protein; neuroprotective; cytoskeletal;
 KW gastrointestinal; immunomodulatory; muscular active; vulnerrary;
 KW gastrointestinal; nephrotropic; antinefactive; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.
 XX
 OS Homo sapiens.
 OS
 PN W0200055180-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05918.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GRNOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Ruben SM;
 DR WPI: 2000-587514/55.
 DR N-PSDB; AAF18318.
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 PS Claim 11; Page 1310-1313; 1425pp; English.
 XX
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytoskeletal; cardioactive;
 CC immunomodulatory; muscular active general; vulnerrary; gastrointestinal
 CC general; nephrotropic; antinefactive; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,

gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders. The proteins may also be used in the treatment of wounds and infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB5549 are used in the course of the invention for the identification and characterisation of the polynucleotide and protein sequences.

SQ Sequence 793 AA;

Query Match	99.48	Score 4006	DB 21	Length 793
Best Local Similarity	99.48	Pred. No. 0		
Matches /65	Conservative 4	Mismatches 1	Indels 0	Gaps 0

Oy	1	MAOMNLOLOJLTPRLKOLJHOLYUPTFPMELROFLAPMIEESODMAYVAAKSESHATVPHNL	60
Dp	24	maqnmqlqldltrylreqhmqlydsfmeelrqflapwiesgdwaaskeeshatlvlhnl	83
Oy	61	LGEIDDOYSRFLQESBNVLQYHNLRRIKOFLQSRYLEKPMELIARIWANCLMEESRLQTA	120
Dp	84	lgeidqgsrflqesbnvlqynhlrrlrlqflqsrylekpmelariwancleesrllqta	143
Oy	121	TAAOOGGAGNPFPAAYVTERKOOMELOHLODYRRKVVQOLEOKMKVVENLODDFERNYTLK	180
Dp	144	taaqgggagmnpitaavterkqgmleqnlqdrtrkvvqoleqkmkvenlqddfdnnytlk	203
Oy	181	SOQDMQDLNNGNOSVTRKMOOLEBMLTALDOMRRISIVSELAGLLSAMEYVQKTUDEEL	240
Dp	204	sqgdmqdlngnngsvtrqkmqlqegmltalqdmrrisivselagllsameyvkltdeel	266
Oy	241	ADMKRRPFIACIGGPNNCLDRLENMTSLAESLOLOROOIKKLEELOOKVSVKGDPIYO	300
Dp	264	admkrtrqfiactggpnncldrleamtlslesqlqtrqrkileeqkvsvkgdpivq	323
Oy	301	HRPMLREBIVELFRNLKMSAFVVEROPCOMPHEDRPLVITGVOFTTKVRLLKVPPELNY	360
Dp	324	hrpmlreelivelfrnlmsafverqpcmhphdrplvltkgvfttkvrllvkpelpny	383
Oy	361	OLKIVICV-IDKXSGDVAALRGSRKNNIIGTNTKVVNMNESNNNGSLSAEFKHLTLEOQSGN	420
Dp	384	qlkikvcldkxsgdvaalrgsrknnlqntkvvnmneesnmngslsaefkhltlreqcgn	443
Oy	421	GGRANCDSLIVTEBELHITFETEVYHQGLIDLETHSLPVVVISNICOMPMAWASILMY	480
Dp	444	ggrancdasilvtebelhltfetevyhqglikdletshlpvvvisnicompmawasilly	503
Oy	481	NNLTNNPNVNFPRPGTMDQDAEVLISWQFSSTTRGSLIDOLITLAKKLGPGGYNYS	540
Dp	504	nmltnpknvnlftkprpigtwdqvaevlswqfssctkrglsleqlltlaaelilppgynys	563
Oy	541	GCQITWAFCKENNAKGFSFWMLDNIIDLKVVYIALALNMEGYIMGFIISKERERALLST	600
Dp	564	gcqitwafckennaqgfsfwvldnildlvkyyialalwnegyimgfiskereralist	623
Oy	601	KPPGFEILRESEESKEGVTFTWEXKIOISGTOQSVNEPTRKQOLNNKMSFAETIIMGKIM	660
Dp	624	kppgfeilrisessekegvltftwexkldisgvtqsvnepytkqqlnmstfaelimgykim	683
Oy	661	DATNIIIVSLPIVLYLPDIPKEAEFAKCYRPSQSEHPEDAPGSAAPYELTKFKICVPTPGSN	720
Dp	684	datnllivslpyllypdipkeeaefkcyrpsqsehpheadpgsaapyltkfkicvptpcsn	743
Oy	721	TIDPMSRTRIDSLMOGCGNNGEGAPBAGOGFESLTITDMDLITSCATSPM 770	
Dp	744	tidpmsrtridslmogcgnngegapbagogfessltitdmetlscatcspm 793	

RESULT	7
AAB19964	
ID	AAB19964 standard; Protein; 770 AA

DT	28-MAR-2001 (first entry)
XX	
DE	Human signal transducer and activator of transcription STAT-3
XX	
KW	STAT-3: signal transducer and activator of transcription 3:
KW	human; crystal; drug screening.

OS Homo sapiens

FT	Key	location/qualifiers
FT	Domain	134...320
FT		/note= "coiled-coil domain"
FT	Domain	321...493
FT		/note= "DNA-binding domain"
FT	Domain	494...583
FT		/note= "linker domain"
FT	Domain	584...688
FT		/note= "SH2 domain"
FT	Region	689...717
FT		/note= "C-terminal tail segment"
FT	Modified-site	705
FT		/note= "O-phosphorylated"

PN	US6160092-A.	
XX		
PD	12-DEC-2000.	
XX		
PF	29-MAY-1998;	98US-0087465
XX		
PR	29-MAY-1998;	98US-0087465

PA (UVRQ) UNIV ROCKEFELLER.

PI Chen X, Darrell JE, Kuriyan J, Vinkemeier U, Zhao Y, Jeruzalmi D
XX
DR WPI: 2001-101568/11.
DR N-PSDB; AAA89229.

PT Novel crystal useful in drug screening assays, comprises portion of
PT signal transducer, activator of transcription and duplex DNA -
XX
PS Disclosure; Column 67-71; 206pp; English.

The present sequence is that of human signal transducer and activator of transcription 3 (STAT-3). The invention provides a crystal of a core portion of a STAT protein in dimer form with an 18-mer duplex DNA (see AaA89233) that contains a binding site for the STAT dimer. The core portion comprises a coiled-coil domain comprising 4 long helices, a DNA binding domain which contains an immunoglobulin-like fold, a C-terminal SH2 domain and a domain that links the DNA binding and SH2 domains. The crystal is of sufficient quality to perform X-ray crystallography studies. Methods of preparing the crystals are included in the invention. Knowledge of the STAT protein's 3-dimensional structure will aid in structure-based drug design. The crystal can be used in drug screening assays to CC identify agonist and antagonist compounds. Antagonists can be used CC to treat inflammation, allergy, asthma and leukaemia, and agonists CC to treat anaemia, neutropenia, thrombocytopenia, cancer, obesity, CC viral diseases, growth retardation, and other conditions CC characterized by insufficient STAT activity.

Sequence 770 AA;

Query Match	98.7%	Score 3977	DB 22	Length 770
Best Local Similarity	98.6%	Pred. No. 0		
Matches 759	Conservative	6	Mismatches 5	Indels 0
				Gaps 0

Oy 1 MAONNLOOLPTRYKJOLHOLYSPFPEMLRQFLAPWIESODMAYAAASKESHATLYVHNL 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 maqwnlqqldtryleqlhqlqsdsfpmeirlqlapwiesdqwayaaskeshatlyvhn1 60

Oy 61 LGELIDQYSRFLOESNNVYQHNLRRIKOFLOSRYLEKPMETARIVACGLWEESRLQTAA 120

|||||
Db 61 lgeIdqysrIqgesnvlyqhnIrrIkqIqIsryIekpmeIarIvarCIweesrIlIqIaa 120
OY 121 TAAOGGQANHPRTAAVYTEKQOMLEOHLODVRRKRVQDEQKKVVENLODDDFNYKTLK 180
Db 121 taaqg9qanhpRTAAVTEKqgmIeqhIdvrtkrvqdlEqkmkvenIqdddfnykclK 180
OY 181 SGGMDQNLGNNGSVTRQKMOOLEOMLTALDQMRRSIVSELAGLLSAMEYOKTLTDEEL 240
Db 181 sqgmdqnlgnngsvtrqkmqIeqmItaldqmrRrsIvseIagllsameyvkltldeei 240
OY 241 ADMKRREPIACIGCPNCLDRLEWMTSLAESQLOTRQOIKKLELOOKSYKCDPIVQ 300
Db 241 adwkrtrqIacIggppnIdrlRnwItslaesqIqtrqIkkleelngkvsykgdplvq 300
OY 301 HRPMLREIVELFRNLMSAFVVERQPCMPHDPRLVIKTCVQPTTKVRLIVKPEELNY 360
Db 301 hrpmlreIvelfrnImksafverqpcmpmhdpdrIvIkTcvqpttkvrlIvKfpeIlny 360
OY 361 QLKIKVCIDKDSGVAALRGSRKFNIIGTNTKVMNMEESNNGLSAERKHLTLRBQRCGN 420
Db 361 qIkIkvcIdkdsqdvaaIrgsrkfInIgtntckvmmeesngsIsaeIkhltlIreqrcgn 420
OY 421 GGRANCOASLIVTEELHITFETEVYHOGKIDLETHSLPVVVISNICOMPMAASILMY 480
Db 421 ggrancdasilvteelhItfeteVyhgIkIdlethslsvvIsnIcQmpnawasilwy 480
OY 481 NMLTNPNKVNVEFTKPIGTDQVAEVLWQFSSPTTKRGLSTIEOLTTLAEKLLGGVNVNS 540
Db 481 nmltnpnkvnvftkprIgtDqvaevlswqfsspttkrglStIeqlltlaekllIpgvnyvs 540
OY 541 GCOITWAKFCKENMAGKGFSPWMLDNIIDLVKKYIILALMNGYIMGFTSKRERAILST 600
Db 541 gcoItwanfckemagkGfsyvwvIdnIdlvkkyIlaImegYImgfIskeraIst 600
OY 601 KPPGFTLLRFESSESKEGCVTFTWEEKDISGKTQIOSVEPYTKOOLNNNSFAEIIIMGYKIM 660
Db 601 kppgftllrfessskEGcvTftweEKdisgktqIosvEpytkqIlnnnsfaeIImgykIm 660
OY 661 DATNILLVSPLYLYLPDIKPEAFGKYCRPESOEHPHEADPGSAAPYLKTKFICVPTTCSN 720
Db 661 datnIlIsplylylpdIpkeefgkyCrpesoeHpeadpgsaapylkKtkfIcvtptcsn 720
OY 721 TIDLPMSPRTLDSLMQSGNNGEAPSPAGGOFESLTFPMDLTSSECATSPM 770
Db 721 tIdlpmSprtlDsImqfGnngEapSagqGfesItfIdmeltsecatSPm 770

RESULT 8
AAR82993
ID AAR82993 standard; protein: 770 AA.
XX
AC AAR82993;
XX
DT 25-MAR-1996 (first entry)
XX
DE Human placenta acute phase response factor protein.
XX
KW human; acute phase response factor; transcription factor;
KW interleukin-6; signal transmission; placenta; antibody; antise; se;
KW ribozyme; antiinflammatory; antitumor; hypotensive; therapy.
XX
OS Homo sapiens.
XX
PN EP676469-A2.
XX
PD 11-OCT-1995.
XX
PF 29-MAR-1995; 95EP-0104670.
XX
PR 04-APR-1994; 94JP-0065825.
XX
PA (KISH/) KISHIMOTO T.

XX Akira S, Kishimoto T;
PI
XX
DR WPI: 1995-346089/45.
DR N-PSDB: AAT05616.
XX
PT New acute phase response factor - for developing inhibitory agents
PT for treating diseases induced by cytokine(s) such as IL-6, e.g.
PT inflammatory diseases
XX
PS Claim 3; Page 9-12; 31pp; English.
XX
CC The sequence corresponds to a human acute phase response factor
CC (APRF), a transcription factor related to signal transmission of
CC interleukin-6 (IL-6). The protein is expressed from a human
CC placenta cDNA. Isolated using an IL-6-treated mouse liver cDNA
CC probe. APRF-inhibitors, e.g. antibodies, antisense
CC oligonucleotides or ribozymes, may be used to treat diseases
CC induced by IL-6, e.g. inflammatory disease, leukemia, cancer,
CC osteoclastia, pulmonary hypertension, etc.
XX
SQ Sequence 770 AA:

Query Match 98.5%; Score 3970; DB 16; length 770;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 758; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
OY 1 MAQNWQLOQDLDRYLKQHLQYSDFFPMELRGFLAPWIESQWATAASKESHATLVFNL 60
Db 1 maqwnqIqldrYleqIhqlYsdsfpmelrqlapwiesqdwyaaskeshatlvfnl 60
OY 61 IGEIDQOVSRLFQESNVLYQOHLRLIKQFLOSRYLEKPMELARIYARCLMEESRLIQTAA 120
Db 61 IgeI IYsrllqesnvlyqhnIrrIkqIqIsryIekpmeIarIvarCIweesrIlIqIaa 120
OY 121 TAAOGGQANHPRTAAVYTEKQOMLEOHLODVRRKRVQDEQKKVVENLODDDFNYKTLK 180
Db 121 taaqg9qanhpRTAAVTEKqgmIeqhIdvrtkrvqdlEqkmkvenIqdddfnykclK 180
OY 181 SGGMDQNLGNNGSVTRQKMOOLEOMLTALDQMRRSIVSELAGLLSAMEYOKTLTDEEL 240
Db 181 sqgmdqnlgnngsvtrqkmqIeqmItaldqmrRrsIvseIagllsameyvkltldeei 240
OY 241 ADMKRREPIACIGCPNCLDRLEWMTSLAESQLOTRQOIKKLELOOKSYKCDPIVQ 300
Db 241 adwkrtrqIacIggppnIdrlRnwItslaesqIqtrqIkkleelngkvsykgdplvq 300
OY 301 HRPMLREIVELFRNLMSAFVVERQPCMPHDPRLVIKTCVQPTTKVRLIVKPEELNY 360
Db 301 hrpmlreIvelfrnImksafverqpcmpmhdpdrIvIkTcvqpttkvrlIvKfpeIlny 360
OY 361 QLKIKVCIDKDSGVAALRGSRKFNIIGTNTKVMNMEESNNGLSAERKHLTLRBQRCGN 420
Db 361 qIkIkvcIdkdsqdvaaIrgsrkfInIgtntckvmmeesngsIsaeIkhltlIreqrcgn 420
OY 421 GGRANCOASLIVTEELHITFETEVYHOGKIDLETHSLPVVVISNICOMPMAASILMY 480
Db 421 ggrancdasilvteelhItfeteVyhgIkIdlethslsvvIsnIcQmpnawasilwy 480
OY 481 NMLTNPNKVNVEFTKPIGTDQVAEVLWQFSSPTTKRGLSTIEOLTTLAEKLLGGVNVNS 540
Db 481 nmltnpnkvnvftkprIgtDqvaevlswqfsspttkrglStIeqlltlaekllIpgvnyvs 540
OY 541 GCOITWAKFCKENMAGKGFSPWMLDNIIDLVKKYIILALMNGYIMGFTSKRERAILST 600
Db 541 gcoItwanfckemagkGfsyvwvIdnIdlvkkyIlaImegYImgfIskeraIst 600
OY 601 KPPGFTLLRFESSESKEGCVTFTWEEKDISGKTQIOSVEPYTKOOLNNNSFAEIIIMGYKIM 660
Db 601 kppgftllrfessskEGcvTftweEKdisgktqIosvEpytkqIlnnnsfaeIImgykIm 660
OY 661 DATNILLVSPLYLYLPDIKPEAFGKYCRPESOEHPHEADPGSAAPYLKTKFICVPTTCSN 720

```

Db      661 datnlllpllylypdipkeeafigkrcpsegehpdeadpgsaapytktkficvptlcsn 720
Oy      721 TTDLMSPPTDLSLMOPGNNGEAPSGAGCFESLTFDMDLTSECATSPM 770
        |||
Db      721 tldlpmpralidslmqnggeapsgagflesltfmdltsecatspm 770

RESULT  9
AA72842
ID      AAY72842 standard; protein: 393 AA.
AC      AAY72842;
XX
DT      31-MAY-2001 (first entry)
XX
DE      Mouse Stat3 protein fragment #3 (378-770 amino acids).
XX
KM      Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
        cellular transformation; dysproliferative disease; cancer; psoriasis;
        therapy.
KM      Mus musculus.
OS      Mus musculus.
PN      WO200116605-A2.
XX
PD      08-MAR-2001.
XX
PF      30-AUG-2000; 2000WO-US23822.
XX
PR      31-AUG-1999; 99US-0387418.
XX
PA      (UYRO ) UNIV ROCKEFELLER.
XX
PI      Zhang X, Horvath C, Wrzeszczynska MH, Darnell JE;
        WPI: 2001-226705/23.
XX
PT      Identifying an agent for use in modulating the interaction between
        transcription factor c-Jun and a Stat3 protein -
        Example 2; Page 68-69; 86pp; English.
XX
TS      The present sequence is mouse Stat3 protein fragment containing 378-770
        amino acids of Stat3 protein. This Stat3 fragment showed very weak
        binding to c-Jun protein in the cell extract.
XX
CC      The invention relates to methods for identifying interacting regions of
        transcription factors and methods for identifying agents which modulates
        the interaction between a transcription factor such as c-Jun and a Stat
        protein such as Stat-1 and Stat-3, useful for modulating gene
        transcription e.g., cellular transformation. These identifying agents are
        used in the treatment of dysproliferative diseases and also for treating
        cancer and psoriasis. A Stat protein comprises the N-terminal domain,
        coiled-coil domain, DNA binding domain, linker domain, SH2 domain and
        transactivation domain.
CC      transactivation domain.
XX
SQ      Sequence 393 AA:

Query Match      51.8%; Score 2087; DB 22; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.5e-172;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      378 LRGSRRFNILGNTKVMNNEESNGSLSAEFKHLTLREORCGNGRANCDA SLVTEELH 437
        |||
Db      1 lrgsrkfnllgntkvmneesngslsaefkhltlreqrcnggrancdaslvtelheh 60
Oy      438 LITFEVYHQGLKIDLEHSLPVYVISNICQPMANASILMYNMLTNPKNVNFTKPP 497
        |||
Db      61 litfevyhgykikidlehslpvvisnicqpmawasilwvnmlltnpkvnftkpp 120
Oy      498 IGTWQVAEVLSSQFSSTTKRGSLISQLTLLAEKLLGPGVNSGCOITNAKCEKEMAGK 557
        |||

```

```

Db      121 lgtwdvvaev]swgfssctkrgisicqlltlaeckllgpgvynsgqjlwakkckenmagk 180
Oy      558 GFSEFWALDNIIIDLVKKYIILALWNEGYIMGFISKERERAILSTRPGTFLLRFSSESKEG 617
        |||
Db      181 gfsfwwlndiildlvkkyilalwnegyimgfiskereralistkpgtctllrfseskeg 240
Oy      618 GVTFTWEXDKISGKTIOISVEPYTKOOLNNSPAEIIIMGYKIMDATNILLVSLVYLYPDI 677
        |||
Db      241 gvtftwexdkisgktiqisvpeyltkqglnmnsfaelimgykimdattllvslplylpydl 300
Oy      678 PKEEAFGKYCRPESGHPPEADPGSAAPYLKTKFICVPTTCSTNTIDLPMSPRTLUSIMOF 737
        |||
Db      301 pkeeafigkrcpsegehpdeadpgsaapytktkficvptlcsnldlpmpralidslmqf 360
Oy      738 GNGEGAPSGAGCFESLTFDMDLTSECATSPM 770
        |||
Db      361 gnngeapsgagflesltfmdltsecatspm 393

RESULT 10
AAR72078
ID      AAR72078 standard; Protein: 750 AA.
AC      AAR72078;
XX
DT      27-SEP-1995 (first entry)
XX
DE      Human Stat91.
XX
KM      Signal transducer and activator of transcription; ISGF-3; STAT;
        Stat91; receptor recognition factor; transcription factor;
        cellular debilitation; derangement; dysfunction; interferon-alpha;
        interferon-gamma.
XX
OS      Homo sapiens.
XX
PN      WO9508629-A.
XX
PD      30-MAR-1995.
XX
PF      26-SEP-1994; 94WO-US10849.
XX
PR      24-SEP-1993; 93US-0126588.
        PR      24-SEP-1993; 93US-0126595.
        PR      11-MAR-1994; 94US-0212184.
        PR      11-MAR-1994; 94US-0212185.
XX
PA      (UYRO ) UNIV ROCKEFELLER.
XX
PI      Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;
        WPI: 1995-139598/18.
XX
PT      N-PSDB; AA089335.
XX
TS      Receptor recognition factor implicated in transcriptional
        stimulation of genes - useful in drug screening assays and/or
        for treating cellular debilitations, derangements and/or
        dysfunctions, etc.
XX
CC      Disclosure: Page 84-88; 160pp; English.
XX
CC      The sequences of cDNA encoding receptor recognition factors having
        mol.wt. of 113 kDa (Stat113), 91 kDa (Stat92) and 84 kDa (Stat84)
        are given in AA089335-37 and the deduced amino acid sequences of the
        STAT proteins in AAR72077-79. These ISGF-3-derived proteins are
        activated by binding of interferon-alpha (all 3 Stat proteins) or
        interferon-gamma (Stat91) to cell receptors.
XX
SQ      Sequence 750 AA:

Query Match      50.1%; Score 2017; DB 16; Length 750;
Best Local Similarity 52.5%; Pred. No. 1.2e-165;

```

Matches 389; Conservative 145; Mismatches 193; Indels 14; Gaps 9;				
OY	1	MAQWNOLODLTRYKLOHLYSDTFPMELOFLAPWIESODMAYAAKSKESHATLVFHNL	60	
Db	1	msqwyelqidskrlqyhyldsfpmeitrylaqlwlekqwehaandsafatirfndl	60	
OY	61	LGEIDQOYSRFLQESNVLYOHNLRIRIKOFLQSRYLEKRMETARIIVARCLMEESRLDTAA	120	
Db	61	lsqiddysrflsenfllqhnirsksrnlqndfndepdqmismlyscskeerkllemaq	120	
OY	121	TAQOQGOAHNPRTAAVYTEKQOMLEOHLODYRKRVQODEQKMKVVENQDDPFDNFYKTLK	180	
Db	121	rftq--aqsgnlgstvmldkqkeldskvrnydkwmclehelkstedqdeydfckll- 177		
OY	181	SQGMODLNGNOSVTRQOKMOOLEOMLTALDOMRSIYSELAGLISAMEYVQKTLTDEEL	240	
Db	178	-gnrehetnyaksdqkqeglllkmylmldnkrkevnhkjiellnveltcgnalindel	236	
OY	241	ADMKRREIACIGGPRNICDLRENMITSLSAESQLQTRQOIKLEELQOKVSYGDPPIVQ	300	
Db	237	vwktrqgsacigppnaciqlqmwftivaeslqvvrqiklleeqkytyehdptlk	296	
OY	301	HRPMLERIVLEFRNLKMSAFVVEROPCMRPHDRPLVITKGVOFTTRVRLVFPRELNY	360	
Db	297	nkqylwdrtfslfqglqgsstfverqpcmrphqprplylktvgvftvklrlvlkqelny	356	
OY	361	OLKIKVCIDKSGVAAALRGSRKFNILOTNTKVMNMESNNGSLSAERKHLTLREQRCGN	420	
Db	357	nlkxkvldfckvnerntvkgfrfnllgthtkmmesetngslaaefrlqikeqk--n	414	
OY	421	GGRANCDASLIVTEBLHLITETEYVHOGKIDLETSHLPVVVISNICOMPANASILWY	480	
Db	415	agrtneqplivteelshsfetqlcpgjvldletslprvvlsvnsqglpsyaasilwy	474	
OY	481	NMLNNPKNVNEFFRKPGTMDQVAEVLNMOFSSFTKRGSLTLEOLTLEAKELGPGVYVS	540	
Db	475	nmlvaepnslfllpparwaqlsevlswqfsyckrlgvndqimlgeklilgpnaspd	534	
OY	541	GCQTWAKFCCKENMAGKFSFWMLNDIIDLVKYIALANEGYIMGPISEKEREALUST	600	
Db	535	gl-1pwtfckekendkfmfwieslleiikhlpilwdngcimgfiskereraallk	593	
OY	601	KPRGTFLLRESSESKGEGVTFTWWEKDIS-GKTQIOSVEPTKQOOLNMMSAEIIMGKI	659	
Db	594	qgpgtfillrisesregaitftweresqnggepdfhavepylkelasvtfpdlirnykv	653	
OY	660	MDANILIVSPLYVLPDIPKEAEFGK-YGRP-ESOEHPDAPGSAAPYLKTKFICVT---	714	
Db	654	maeenlpenplkylpnlidkdhafgkystrpkeapeimeldgprkgtgylkellsvsevh	713	
OY	715	PTTCSNITD-LPMSPRITLDSL 734		
Db	714	psrlqtdnllpmspeefdev 734		
RESULT 11				
AAW03168				
ID	AAW03168 standard; Protein; 750 AA.			
AC	AAW03168:			
XX	24-OCT-1996 (first entry)			
DT	Human STAT1-alpha.			
XX	STAT: STAT1-alpha; signal transducer and activator of transcription;			
KM	DNA binding protein; ligand: receptor; oncogenes; inflammation;			
KM	autoimmune disease; antagonist; therapy; STAT1.			
XX	Homo sapiens.			
OS	Homo sapiens.			
XX	Key			
XX	FH			
FT	Domain			
FT	Location/Qualifiers			
FT	399..508			

FT	/label= DNA-binding-domain
FT	/note= Claim 3, page 110
FT	Modified-site
FT	701
FT	/label= Phosphorylation_site
FT	Modified-site
FT	727
FT	/label= Phosphorylation-site
XX	
PN	MO9620954-A2.
XX	
PD	11-JUL-1996.
XX	
PF	28-DEC-1995: 95MO-US17025.
XX	
PR	06-JAN-1995: 95US-0369796.
XX	
PA	(UYRQ) UNIV ROCKEFELLER.
XX	
PI	Darnell JF, Horvath CM, Wen Z, Zhong Z;
XX	
XX	WPI: 1996-333941/33.
DR	N-PSDB: AAT 476.
XX	
PT	New STAT protein DNA-binding domain peptide(s) - useful for
PT	diagnosing, preventing or treating cellular dysfunction, e.g.
PT	oncogenes, inflammation, parasitic disease or autoimmunity
XX	
PS	Disclosure; Page 69-73; 138pp; English.
XX	
CC	Signal transducer and activator of transcription (STAT) protein
CC	STAT1-alpha (AAW03168), also known as STAT1, is a 91 kDa protein
CC	having a dual purpose, i.e. signal transduction from ligand-activated
CC	receptor kinase complexes followed by nuclear translocation and DNA
CC	binding to activate transcription. Recombinant STAT1-alpha can
CC	be obtd. using an isolated cDNA clone (AAT31276). STAT1-alpha
CC	includes a DNA-binding domain (see also AAW03165) capable of both
CC	receptor recognition and message delivery via DNA binding in a
CC	receptor-ligand specific manner. Maximum STAT1-alpha activation
CC	of genes requires phosphorylation of both Tyr-701 and Ser-727.
CC	STAT proteins and their DNA binding domains (see also AAW03165-67,
CC	AAW03169-76) are useful for screening antagonists used to inhibit
CC	STAT-mediated signal transduction and activation of transcription.
XX	
SO	Sequence 750 AA:
Query Match	
Best Local Similarity 52.5%; Pred. No. 1,2e-165;	
Matches 389; Conservative 145; Mismatches 193; Indels 14; Gaps 9;	
OY	1 MAQWNOLODLTRYKLOHLYSDTFPMELOFLAPWIESODMAYAAKSKESHATLVFHNL
Db	1 msqwyelqidskrlqyhyldsfpmeitrylaqlwlekqwehaandsafatirfndl
OY	61 LGEIDQOYSRFLQESNVLYOHNLRIRIKOFLQSRYLEKRMETARIIVARCLMEESRLDTAA
Db	61 lsqiddysrflsenfllqhnirsksrnlqndfndepdqmismlyscskeerkllemaq
OY	121 TAQOQGOAHNPRTAAVYTEKQOMLEOHLODYRKRVQODEQKMKVVENQDDPFDNFYKTLK
Db	121 rftq--aqsgnlgstvmldkqkeldskvrnydkwmclehelkstedqdeydfckll- 177
OY	181 SQGMODLNGNOSVTRQOKMOOLEOMLTALDOMRSIYSELAGLISAMEYVQKTLTDEEL
Db	178 -gnrehetnyaksdqkqeglllkmylmldnkrkevnhkjiellnveltcgnalindel
OY	241 ADMKRREIACIGGPRNICDLRENMITSLSAESQLQTRQOIKLEELQOKVSYGDPPIVQ
Db	237 vwktqgsacigppnaciqlqmwftivaeslqvvrqiklleeqkytyehdptlk
OY	301 HRPMLERIVLEFRNLKMSAFVVEROPCMRPHDRPLVITKGVOFTTRVRLVFPRELNY
Db	297 nkqylwdrtfslfqglqgsstfverqpcmrphqprplylktvgvftvklrlvlkqelny

DR N-PSDB: AA089337.

XX Receptor recognition factor implicated in transcriptional
PT stimulation of genes - useful in drug screening assays and/or
PT for treating cellular debilitations, derangements and/or
PT dysfunctions, etc.

PS Disclosure: Page 91-94; 160pp; English.

XX The sequences of cDNA encoding receptor recognition factors having
CC mol. wt. of 113 kDa (Stat113), 91 kDa (Stat92) and 84 kDa (Stat84)
CC are given in AA089335-37 and the deduced amino acid sequences of the
CC STAT proteins in AAR7207-79. These ISGF-3-derived proteins are
CC activated by binding of Interferon-alpha (all 3 Stat proteins) or
CC Interferon-gamma (Stat91) to cell receptors.

XX Sequence 712 AA:

Query Match 49.68; Score 1997; DB 16; Length 712;

Best Local Similarity 53.18; Pred. No. 5.8e-164;

Matches 381; Conservative 142; Mismatches 184; Indels 10; Gaps 7;

QY 1 MAONNLOQLDTRYLQHLQSLSDTFPMELROFLAPMTESODMAYAAKSHATLVFNL 60
D 1 msqyvelqiskflegbqlyddsfmeirgylaqwlekqdehaavsfatirhnl 60
QY 61 LGEIDDOYSRFLQESNVLQHNLRIRKQFLOSRYLEKPMETARIVAKCLMEESRLLOTAA 120
D 61 lsgiddqysrfslenfllqhnirskrnldnfgedpigmmllyscleerki1lenag 120
QY 121 TAAOGGAGNHPRTAAVYVEKQOMLEQHLQDVRKRVQDLEOKMKYVENLODDDFNKTLLK 180
D 121 rfnq--agsgniqstwmldkqelaskvrvnkdkwmclelksiedlqdeydfkckcl- 177
QY 181 SQGMQDNGNNGSVTRQKMOQLBOMLTALDOMRSIVSELAGLSAMEYOKLTDEEL 240
D 178 -qnrehetnyaksdqgeqlilkmylmdnkrkevvhkiellnvtelqnalindel 236
QY 241 ADKKRREINACIGPPNCLDRLENWITSLAESLOLQROIKLEEDOOKVSYKGDPIVQ 300
D 237 vewrttqgsaciqppnacdqlgnwflvaseiqvvrqklkieleeqkytyehndpitk 296
QY 301 HRPMLEERIVELFRNLKMSAFVVEROPCMRPHRPLVTKGVOPFTVRLLVKKPELNY 360
D 297 nkqylwdrtslfgqllqgsfverqpcmhprqprlylktgvqftvkirllvklqelny 356
QY 361 QLKTKVICIDSGVAAALRGSRKFNILGTNTKVMNMEESNNGSLSAEPKHLTLREORCN 420
D 357 nlkxkvlfcdvnerntvkgfrkfnlglthkvmmeestngslaaefrhqlkeqk--n 414
QY 421 GGRANCDASLIVTEELHLPETEVEYHOGKIDLETHSLPVVTSNLCOMNANASILMY 480
D 415 agltneqplivteelnsistecqicqpglvidletcsiprvvisnvsgqlpsgwasllwy 474
QY 481 NMLTNPNKNVNFETKPPDIGTMDOVAEVLWQFSSTTRKGLSIEOLTLAEKLLGPGVNS 540
D 475 nmivaepnrlsfltpccarwaqjsevlswqfsstkrqlnvqdqlmgeklignaspd 534
QY 541 GCQITMAKKPKENAGGFSFWWLDNIIDLKXYILALWNEGYIMGFISKEREHALST 600
D 535 gl-1pwrtckenlndnfjwiesllei1khhllplwngcimgfiskerera1lkd 593
QY 601 KPPGTLLRFSSESKEGVFTWWEKDIS-GKTIQSVPEPTKOQLNMNMFALIIIMGYKI 659
D 594 qpggtflirtseesregalftwversqnggepfhavpytkelsavtfdllrnykv 653
QY 660 MDAITNIVSLVLYLPDIPKEAEFGK-YCRP-ESQEHPEADPGSADPYLKTFTICVT 714
D 654 maenlpenpiklylpyldkdhafgkyrsrpkeapepemdqpkgtgtyiktellsvs 710

Search complete
Job time: 4843 sec
Mar 19, 2002, 15:22:01

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:21:58 ; Search time 56.45 Seconds
(without alignments)
1995.212 Million cell updates/sec

Title: US-08-212-185-12

Perfect score: 4029
Sequence: 1 MAQNNQLQQLDTRFLKQLHQ.....QFESLTFMDLTSQCATSPM 770

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_ricent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4006	99.4	770	4	014916
2	3891.5	99.1	769	4	09BM54
3	3839.5	95.3	769	13	09PVX8
4	3474	86.2	806	13	093599
5	3373.5	83.7	767	13	013133
6	2038	50.6	749	11	090XK0
7	2026	50.3	749	11	090XJ2
8	2006	49.8	712	11	099K94
9	1901	47.2	754	13	013131
10	1891	46.9	754	13	013132
11	1786.5	44.3	749	13	093598
12	1597.5	39.7	394	13	09DDJ8
13	1287	31.9	922	11	090XJ2
14	1284.5	31.9	925	11	090ZE4
15	873	21.7	784	13	09BG74
16	867	21.5	787	13	093378
17	865.5	21.5	786	11	09JRM1
18	834	20.7	763	6	09N145
19	663.5	16.5	592	13	09PWP7

20	566	14.0	679	4	09BOD2	09bq2 homo sapien
21	532	13.2	722	5	097164	097164 anopheles g
22	473.5	11.8	156	4	09UDL5	09ud5 homo sapien
23	423	10.5	703	5	09NAD6	09nad6 caenorhabdl
24	389.5	9.7	195	6	09KRY9	09kry9 sus scrofa
25	342	8.5	155	4	09UDL4	09ud4 homo sapien
26	299.5	7.4	144	11	070428	070428 rattus norv
27	297	7.4	111	11	070406	070406 rattus norv
28	292	7.2	56	6	09N0J4	09n04 bos laurus
29	280.5	7.0	177	6	09N0E5	09n05 bos laurus
30	273	6.8	51	11	099ML3	099ml3 mus musculu
31	266	6.6	51	4	09BXH2	09bxh2 homo sapien
32	256.5	6.4	85	6	029340	029340 sus scrofa
33	251.5	6.2	141	11	070429	070429 rattus norv
34	248.5	6.2	230	11	09R0X9	09r0x9 mus musculu
35	224	5.6	43	11	09QVR4	09qvr4 rattus sp.
36	223.5	5.5	205	11	09R0X8	09r0x8 mus musculu
37	215	5.3	85	6	029356	029356 sus scrofa
38	211.5	5.2	929	5	09BLX2	09blx2 dictyostell
39	178	4.4	43	11	09QVR3	09qvr3 rattus sp.
40	177.5	4.4	707	5	000910	000910 dictyostell
41	164.5	4.1	1356	4	014707	014707 homo sapien
42	161	4.0	2041	4	09HAT7	09hat7 homo sapien
43	161	4.0	2047	4	09UH61	09uh61 homo sapien
44	159.5	4.0	2101	4	014981	014981 homo sapien
45	156.5	3.9	1300	4	013999	013999 homo sapien

ALIGNMENTS

RESULT 1
ID 014916
AC 014916
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-AUG-1999 (TREMBLrel. 11, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE TRANSCRIPTION FACTOR (SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION
DE 3).
GN STAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=9826260; PubMed=9630560.
RA Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.,
RT "Highly conserved amino-acid sequence between murine STAT3 and a
RL Gene 213:119-124(1998).
RN 12
RP SEQUENCE OF 564-704 FROM N.A.
RC TISSUE=HEPATOMA.
RA Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.,
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ012463; CAA10032.1;
DR EMBL: AF029111; AAB84254.1;
DR HSSP: P42227; IBGI.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PSS0001; SH2; 2.
SQ SEQUENCE 770 AA; 88067 MW; 6C00632211C8012D CRC64;

Query Match 99.4%; Score 4006; DB 4; Length 770;
Best Local Similarity 99.4%; Pred. No. 3.2e-263;
Matches 765; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:25:48 ; Search time 16.05 Seconds
(without alignments)
3654.480 Million cell updates/sec

Title: US-08-212-185-12

Sequence: 1 MAQWNOLOQUDTRYLKQLHQ.....QFESLTEDMDLTSCCATSPM 770

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : PIR68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	442	57.4	770	2	I49508	ISGF3 p91-related
2	221	28.7	770	2	A54444	DNA-binding protein
3	12	1.6	739	2	A46159	Interferon-dependent
4	11	1.4	57	2	S63680	signal transducer
5	11	1.4	65	2	S63679	signal transducer
6	11	1.4	748	2	A56047	gamma-interferon a
7	11	1.4	851	2	A46160	interferon alpha-1
8	9	1.2	44	2	T26893	hypothetical prote
9	9	1.2	72	2	T26170	hypothetical prote
10	8	1.0	293	2	G82180	transcription regu
11	8	1.0	663	2	S67259	transcription regu
12	8	1.0	728	2	F84279	amino acid transp
13	8	1.0	786	2	I49274	mammary gland fact
14	8	1.0	793	2	S54772	mammary gland fact
15	8	1.0	794	2	G02317	transcription acti
16	8	1.0	794	2	S55527	mammary g-and fact
17	8	1.0	837	2	I57557	DNA-binding protei
18	8	1.0	848	2	A54740	interleukin-4-indu
19	8	1.0	1088	1	P1XRBR	inner layer protei
20	8	1.0	1088	1	P1XRBR	inner layer protei
21	8	1.0	1088	1	P1XRBR	inner layer protei
22	8	1.0	1088	2	S39261	VP1 protein - porc
23	8	1.0	1088	2	S13558	VP1 protein - bov
24	7	0.9	49	2	S63682	signal transducer
25	7	0.9	64	2	S63681	signal transducer
26	7	0.9	77	2	H70642	probable ribosomal
27	7	0.9	94	2	F31844	spda protein - Str
28	7	0.9	111	2	T17132	hypothetical prote
29	7	0.9	128	2	H70435	hypothetical prote

ALIGNMENTS

RESULT	1	ALIGNMENTS
I49508	ISGF3 p91-related transcription factor - mouse	
C/Species: Mus musculus (house mouse)		
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000		
C/Accession: I49508; I49009		
R/Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.;		
Cell 77, 63-71, 1994		
A>Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related		
A/Reference number: A54444, PMID:94208062		
A/Accession: I49508		
A>Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: mRNA		
A:Residues: 1-770 <RES>		
A:Cross-references: GB:I29278; NID:q476715; PIDN:AAA37254.1; PID:q476716		
R/Raz, R.; Durbin, J.E.; Levy, D.E.		
J. Biol. Chem. 269, 24391-24395, 1994		
A>Title: Acute phase response factor and additional members of the interferon-stimula		
A/Reference number: I49009; PMID:95014185		
A/Accession: I49009		
A>Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: mRNA		
A:Residues: 1-393, 'M', 395-700, 702-770 <RE2>		
A:Cross-references: EMBL:U08378; NID:q473889; PIDN:AAA56668.1; PID:q473890		
C:Genetics: APRF		
C:Superfamily: human signal transducer and transcription activator STAT5A		
Query Match	57.4%; Score 442; DB 2; Length 770;	
Best Local Similarity	99.6%; Pred. No. 0;	
Matches 742; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
OY	26 FPEMLRQFLAPWIESQDMAYAAKESHATLVFNLLGEIDQYSRFLQESNVLYQHLNR	85
DB	26 FPEMLRQFLAPWIESQDMAYAAKESHATLVFNLLGEIDQYSRFLQESNVLYQHLNR	85
OY	86 IKQFLOSRYLEKPEMETARIYARCLMESRLLOTAATAAGCGGANHPYAAVTEKQOMLE	145
DB	86 IKQFLOSRYLEKPEMETARIYARCLMESRLLOTAATAAGCGGANHPYAAVTEKQOMLE	145
OY	146 QHLDQVKKRVQDLEQKKVVENLDQDDPFYKTLKSGDMQDINGNNOSTYRQKMOOLEQ	205
DB	146 QHLDQVKKRVQDLEQKKVVENLDQDDPFYKTLKSGDMQDINGNNOSTYRQKMOOLEQ	205
OY	206 MLTALDDMRRSIVSELAGLISAMEYVOKTLTDELDADWKRREPETACTGSPNICTDLRLEN	265
DB	206 MLTALDDMRRSIVSELAGLISAMEYVOKTLTDELDADWKRREPETACTGSPNICTDLRLEN	265
OY	266 WITSLSAQSLOTROQRIKKLEELQOKVSKGDPVQHRPMLSEETIVELFRMLMSAVYVR	325
DB	266 WITSLSAQSLOTROQRIKKLEELQOKVSKGDPVQHRPMLSEETIVELFRMLMSAVYVR	325

QY 326 OPCPMHDPRLVYIKTGVQFTTKYRLLYKPELNVOLKIKYCIDKSDGVAALRGRKFN 385
 DB 326 OPCPMHDPRLVYIKTGVQFTTKYRLLYKPELNVOLKIKYCIDKSDGVAALRGRKFN 385
 QY 386 ILGNTNTKVMNNEESNNGSLSAEFKHLTLREORCGNGCRANCASLIYEELHLITFEFEV 445
 DB 386 ILGNTNTKVMNNEESNNGSLSAEFKHLTLREORCGNGCRANCASLIYEELHLITFEFEV 445
 QY 446 YHOGKLTLELHSLPYVYVINSICMPNMAASILWYNNLTNPKVNFPTKPPIGTWDOVA 505
 DB 446 YHOGKLTLELHSLPYVYVINSICMPNMAASILWYNNLTNPKVNFPTKPPIGTWDOVA 505
 QY 506 EVLSMOSSTTKRGLSIEQLTTLAEKLLGPGVNSGCCITNAKCKEMKAGKGSFMYWL 565
 DB 506 EVLSMOSSTTKRGLSIEQLTTLAEKLLGPGVNSGCCITNAKCKEMKAGKGSFMYWL 565
 QY 566 DNIITDYKKYLLALMNEGYIMGFISKERERAILSTKPPGFTLLRPFSSSKGEGVTFWE 625
 DB 566 DNIITDYKKYLLALMNEGYIMGFISKERERAILSTKPPGFTLLRPFSSSKGEGVTFWE 625
 QY 626 KDISGKTQIOSEVEPYTKOQLNNMSFAELIMGYKIMDATNIIIVSPLYLYPIPKREARFK 685
 DB 626 KDISGKTQIOSEVEPYTKOQLNNMSFAELIMGYKIMDATNIIIVSPLYLYPIPKREARFK 685
 QY 686 YCPRESOEHPEADPGSAAPYLKTKFICVPTTCSNTIDLPMSPTLDSLMQFGNGEGAE 745
 DB 686 YCPRESOEHPEADPGSAAPYLKTKFICVPTTCSNTIDLPMSPTLDSLMQFGNGEGAE 745
 QY 746 PSAGGOFESLTFMDLTSECATSPM 770
 DB 746 PSAGGOFESLTFMDLTSECATSPM 770

RESULT 2
 A:Accession: A54444
 C:Species: Homo sapiens (man)
 C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 28-Jul-2000
 C:Accession: A54444
 R:Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.-J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Sud
 Cell 77, 63-71, 1994
 A:Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra
 A:Reference number: A54444; MUID:94208062
 A:Accession: A54444
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-770 <RES>
 A:Cross-references: GB:L29277; NID:9475788; PID:9475789
 C:Genetics:
 A:Gene: GDB:STAT3; APRF
 A:Cross-references: GDB:350950
 A:Map position: 17q21-17q21
 C:Superfamily: human signal transducer and transcription activator STAT5A
 C:Keywords: DNA binding; transcription factor

Query Match 28.7%; Score 221; DB 2; Length 770;
 Best Local Similarity 100.0%; Pred. No. 1.1e-220;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 FPMELRQFLAPWIESQMAAYAAKESHATLVFHNLTGEIDDOYSRPFQESVLYOHNLRR 85
 DB 26 FPMELRQFLAPWIESQMAAYAAKESHATLVFHNLTGEIDDOYSRPFQESVLYOHNLRR 85
 QY 86 IKOFLQSRVLEKPEWETIARIARCIWESRLLQTAATAAGOGGANGHPJAAVTEKQOMLE 145
 DB 86 IKOFLQSRVLEKPEWETIARIARCIWESRLLQTAATAAGOGGANGHPJAAVTEKQOMLE 145
 QY 146 OHLODVAKRVQDLEQKKKVVENLODDDFNYKTLKSGQDMODLNGNOSYTRQKMOOLEQ 205
 DB 146 OHLODVAKRVQDLEQKKKVVENLODDDFNYKTLKSGQDMODLNGNOSYTRQKMOOLEQ 205

QY 206 MTLALDQMRRSIVSELACLSAMEYVOKTITDELDADMKRR 246
 DB 206 MTLALDQMRRSIVSELACLSAMEYVOKTITDELDADMKRR 246
 RESULT 3
 A:Accession: A46159
 C:Species: Homo sapiens (man)
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Jul-2000
 C:Accession: A46159
 R:Schindler, C.; Fu, X.-Y.; Improt, T.; Abersold, R.; Darnell Jr., J.E.
 Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
 A:Title: Proteins of transcription factor ISGF-3: one gene encodes the 91 and 84-kDa
 A:Reference number: A46159; MUID:9236557
 A:Accession: A46159
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid; protein
 A:Residues: 1-739 <SCH>
 A:Experimental source: HeLa cells
 A:Note: sequence extracted from NCI backbone (NCBI:P110818)
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.6%; Score 12; DB 2; Length 739;
 Best Local Similarity 100.0%; Pred. No. 0.0013;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 585 IMGFISKERERA 596
 DB 585 IMGFISKERERA 578

RESULT 4
 S:63680
 Signal transducer and activator of transcription 2 type c - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 02-Sep-2000
 C:Accession: S: 680
 R:Sugiyama, T.; Nishio, Y.; Kishimoto, T.; Akira, S.
 FEBS Lett. 381, 191-194, 1996
 A:Title: Identification of alternative splicing form of Stat2.
 A:Reference number: S:63679; MUID:96176320
 A:Accession: S:63680
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 1-57 <SUG>
 A:Cross-references: GB:S81491; NID:91478435; PIDN:AAB36226.1; PID:9147843;
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.4%; Score 11; DB 2; Length 57;
 Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLRFSSES 614
 DB 1 GTFLRFSSES 11

RESULT 5
 S:63679
 Signal transducer and activator of transcription 2 type a - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 02-Sep-2000
 C:Accession: S:63679
 R:Sugiyama, T.; Nishio, Y.; Kishimoto, T.; Akira, S.
 FEBS Lett. 381, 191-194, 1996
 A:Title: Identification of alternative splicing form of Stat2.
 A:Reference number: S:63679; MUID:96176320
 A:Accession: S:63679
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA

A:Residues: 1-65 <SUG>
 A:Cross-references: GB:S81491; NID:q1478435; PIDN:AA636225.1; PID:q1478436
 C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.4%; Score 11; DB 2; Length 65;
 Best Local Similarity 100.0%; Pred. No. 0.0017;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLRFSS 614
 |||||
 Db 1 GTFLRFSS 11

RESULT 6
 A56047
 gamma-interferon activation site-binding protein Stat4 - mouse

C:Species: Mus musculus (house mouse)
 C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 28-Jul-2000
 C:Accession: A56047
 R:Yamamoto, K.; Quelle, F.W.; Thierfelder, W.E.; Kreider, B.L.; Gilbert, D.J.; Jenkins, M.J. Cell. Biol. 14, 4342-4349, 1994
 A:Title: Stat4, a novel gamma interferon activation site-binding protein expressed in e
 A:Reference number: A56047; MUID:94277038
 A:Accession: A56047
 A>Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-748 <YAM>
 A:Cross-references: GB:U09351; NID:q509502; PIDN:AA19692.1; PID:q509503
 C:Superfamily: human signal transducer and transcription activator STAT5A
 C:Keywords: DNA binding; phosphoprotein

Query Match 1.4%; Score 11; DB 2; Length 748;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 603 PGFLRFSES 613
 |||||
 Db 592 PGFLRFSES 602

RESULT 7
 A46160
 interferon alpha-induced transcription activator ISGF-3, 113K chain - human

N:Alternate names: stat2 protein
 C:Species: Homo sapiens (man)
 C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Aug-2000
 C:Accession: A46160; S71908; S53873
 R:Fu, X.Y.; Schindler, C.; Imbrota, T.; Aebersold, R.; Darnell Jr., J.E. Proc. Natl. Acad. Sci. U.S.A. 89, 7840-7843, 1992
 A:Title: The proteins of ISGF-3, the interferon alpha-induced transcriptional activator,
 A:Reference number: A46160; MUID:92366558
 A:Accession: A46160
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA; protein
 A:Residues: 1-851 <FU>
 A>Note: sequence extracted from NCBI backbone (NCBI:110820)
 R:Yan, R.; Qureshi, S.; Zhong, Z.; Men, Z.; Darnell, J.E. submitted to the EMBL Data Library, December 1994
 A:Reference number: S71908
 A:Accession: S71908
 A:Molecule type: DNA
 A:Residues: 1-851 <YAN>

A:Cross-references: EMBL:U18671; NID:q1293919; PIDN:AA98760.1; PID:q1293920
 R:Yan, R.; Qureshi, S.; Zhong, Z.; Men, Z.; Darnell Jr., J.E. Nucleic Acids Res. 23, 459-463, 1995

A:Title: The genomic structure of the STAT genes: multiple exons in coincident sites in
 A:Reference number: S53873; MUID:95192056
 A:Accession: S53873
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-196;392-591;684-730 <YAN>

A:Cross-references: EMBL:U18671
 C:Genetics:
 A:Gene: stat2
 A:Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2;
 C:Superfamily: human signal transducer and transcription activator STAT5A
 C:Keywords: signal transduction; transcription regulation

Query Match 1.4%; Score 11; DB 2; Length 851;
 Best Local Similarity 100.0%; Pred. No. 0.017;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLRFSS 614
 |||||
 Db 596 GTFLRFSS 606

RESULT 8
 T26893
 hypothetical protein Y44A6D.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T26893
 R:Ainscough, R. submitted to the EMBL Data Library, June 1998
 A:Reference number: Z20282
 A:Accession: T26893
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-44 <WIL>
 A:Cross-references: EMBL:AL023842; PIDN:CAA19514.1; GSPDB:GN00023; CESP:Y44A6D.1
 A:Experimental source: clone Y44A6D
 C:Genetics:
 A:Gene: CESP:Y44A6D.1
 A:Map position:
 A:Introns: 30/1

Query Match 1.2%; Score 9; DB 2; Length 44;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 KLEELQOK 290
 |||||
 Db 33 KLEELQOK 41

RESULT 9
 T26170
 hypothetical protein W04G5.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T26170
 R:Kershaw, J. submitted to the EMBL Data Library, March 1997
 A:Reference number: Z20164
 A:Accession: T26170
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-72 <WIL>
 A:Cross-references: EMBL:Z93391; PIDN:CAB07680.1; GSPDB:GN00019; CESP:W04G5.3
 A:Experimental source: clone W04G5
 C:Genetics:
 A:Gene: CESP:W04G5.3
 A:Map position: 1
 A:Introns: 20/1

Query Match 1.2%; Score 9; DB 2; Length 72;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 KLEELQOK 290

Db 23 KKEELQOK 31

RESULT 10

G82180

transcription regulator LysR family VC1588 [imported] - *Vibrio cholerae* (strain N16961)

C:Species: *Vibrio cholerae*

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: G82180

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Carlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.

I: R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A:Reference number: A82035; MUID:2040683

A:Accession: G82180

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-293 <HEI>

A:Cross-references: GB:AE004236; GB:AE003852; NID:g9656095; PIDN:AAF94742.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1588

A:Map position: 1

C:Superfamily: *Pseudomonas putida* regulatory protein catr

Query Match 1.0%; Score 8; DB 2; Length 293;

Best Local Similarity 100.0%; Pred. No. 8.4;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 QOIKKLEE 286

Db 35 QOIKKLEE 42

RESULT 11

S67259

MNE1 protein - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein 06353; protein YOR350C

C:Species: *Saccharomyces cerevisiae*

C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000

C:Accession: S67259; S67262; S20175; S67404; S19073

R:Goffeau, A.; Purnelle, B.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67246

A:Accession: S67259

A:Molecule type: DNA

A:Residues: 1-663 <GOR>

A:Cross-references: EMBL:275258; NID:g1420761; PID:g1420762; MIPS:YOR350C

A:Experimental source: strain S288C

R:Deilus, H.; Hebling, U.; Hofmann, B.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67261

A:Accession: S67262

A:Molecule type: DNA

A:Residues: 1-663

A:Cross-references: EMBL:275258; NID:g1420761; PID:g1420762; MIPS:YOR350C

A:Experimental source: strain S288C

R:Leem, S.H.; Ogawa, H.

Nucleic Acids Res. 20, 449-457, 1992

A:Title: The MRP4 gene encodes a novel protein kinase homologue required for meiotic recombination

A:Reference number: S20174; MUID:92158649

A:Accession: S20175

A:Molecule type: DNA

A:Residues: 1-219 <LEES>

A:Cross-references: EMBL:X63112; NID:g3968; PIDN:CAA44826.1; PID:g3970

R:Purnelle, B.; Goffeau, A.

submitted to the EMBL Data Library, February 1996

A:Description: Nucleotide sequence analysis of a 40 kb segment on the right arm of yeast

ome 1 genes.

A:Reference number: S67392

A:Accession: S67404

A:Molecule type: DNA

A:Residues: 1-663 <PUR>

A:Cross-references: EMBL:X95720; NID:g1199839; PID:g1199852

C:Genetics:

A:Gene: SCD:MNE1

A:Cross-references: SCD:S0005877; MIPS:YOR350C

A:Map position: 15R

C:Superfamily: *Saccharomyces cerevisiae* MNE1 protein

Query Match 1.0%; Score 8; DB 2; Length 663;

Best Local Similarity 100.0%; Pred. No. 17;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 590 SKERERAI 597

Db 249 SKERERAI 256

RESULT 12

F84279

amino acid transporter [imported] - *Halobacterium* sp. NRC-1

C:Species: *Halobacterium* sp. NRC-1

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: F84279

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky

; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja

Jung, K.H.; Alam, M.; Freltas, F.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Edhardt, H.; Lowe, T.M.

A:Title: Genome sequence of *Halobacterium* species NRC-1.

A:Reference number: A84160; MUID:20504483

A:Accession: F84279

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-728 <STO>

A:Cross-references: GB:AE004437; NID:g10580770; PIDN:AAG19602.1; GSPDB:GN00138

C:Genetics:

A:Gene: ynhg

Query Match 1.0%; Score 8; DB 2; Length 728;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 ELAQLLSA 227

Db 653 ELAQLLSA 660

RESULT 13

I49274

mammary gland factor - mouse

N:Alternate names: STAT5 protein homolog p80

C:Species: *Mus musculus* (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 01-Dec-2000

C:Accession: I49274; S54773; S54727

R:Jin, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.

Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995

A:Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involve

A:Reference number: I49273; MUID:96004632

A:Accession: I49274

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-786 <RES>

A:Cross-references: EMBL:U21110; NID:g747973; PIDN:AAC52282.1; PID:g747974

R:Wu, J.L.F.; Makao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.

EMBO J. 14, 1166-1175, 1995

A:Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleu

A:Reference number: S54772; MUID:95237198

A:Accession: S54773

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA
A:Residues: 1-432, 'E', 434-786 <MU1>
A:Cross-references: EMBL:248539; NID:g758635; PIDN:CAA8420.1; PID:g758636
R:Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Queller, F.; Basu, R.; Saris, E.M.O. J. 14, 1402-1411, 1995
A:Title: Interleukin-3 signals through multiple isoforms of Stat5.
A:Reference number: S54725; MUID:95246733
A:Accession: S54727
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-432, 'E', 434-786 <AZA>
C:Genetics:
A:Gene: Stat5b
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.0%; Score 8; DB 2; Length 786;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
|||||
DB 613 GTFLLRFS 620

RESULT 14
S54772
mammary gland factor - mouse
N:Alternate names: stat5 protein
C:Species: Mus musculus (house mouse)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 28-Jul-2000
C:Accession: S54772; 149273
R:Mu, A.L.F.; Waked, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A:Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A:Reference number: S54772; MUID:95237198
A:Accession: S54772
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-793 <MU1>
A:Cross-references: EMBL:248538; NID:g758633; PIDN:CAA8419.1; PID:g758634
R:Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A:Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved in
A:Reference number: 149273; MUID:96004632
A:Accession: 149273
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-793 <RES>
A:Cross-references: EMBL:U21103; NID:g747971; PIDN:AAA80590.1; PID:g747972
C:Genetics:
A:Gene: Stat5a
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.0%; Score 8; DB 2; Length 793;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
|||||
DB 613 GTFLLRFS 620

RESULT 15
G02317
transcription activator stat5a - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 28-Jul-2000
C:Accession: G02317
R:Liu, J.
submitted to the EMBL Data Library, December 1995
A:Reference number: H01043

A:Accession: G02317
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-794 <LIN>
A:Cross-references: EMBL:U43185; NID:g1151169; PIDN:AAB06589.1; PID:g1151170
C:Superfamily: human signal transducer and transcription activator STAT5A

Query Match 1.0%; Score 8; DB 2; Length 794;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 GTFLLRFS 611
|||||
DB 613 GTFLLRFS 620

Search completed: March 19, 2002, 15:29:01
Job time: 193 s.

```
QY 1 MAOMNLOOQLODTRVLEQJHOLYSDFPMELOKFLAPWIESODMAYAAKESHATLVPHNL 60
D 1 MAOMNLOOQLODTRVLEQJHOLYSDFPMELOKFLAPWIESODMAYAAKESHATLVPHNL 60
QY 61 LGEIDDOQYSRFLQESNNVLYOHNLKRIKQFLOSRYLEKPMELIARIVARCLMEESRLQTA 120
D 61 LGEIDDOQYSRFLQESNNVLYOHNLKRIKQFLOSRYLEKPMELIARIVARCLMEESRLQTA 120
QY 121 TAAOOGGAGNHPTAAVTEKQOOMEQHLQDVYKRVQDLEOKKVVENLODDFENYKTLK 180
D 121 TAAOOGGAGNHPTAAVTEKQOOMEQHLQDVYKRVQDLEOKKVVENLODDFENYKTLK 180
QY 181 SOGDMDLNGNNOSTYRQKMOOLEQMLTALDOMRSIYSELAGLLSAMEYVQKTLTDEEL 240
D 181 SOGDMDLNGNNOSTYRQKMOOLEQMLTALDOMRSIYSELAGLLSAMEYVQKTLTDEEL 240
QY 241 ADMKRREPETIACIGPPNICDLRENNITSLAESQLOTRQOIKKLEELQOKVSYKGDPIVQ 300
D 241 ADMKRREPETIACIGPPNICDLRENNITSLAESQLOTRQOIKKLEELQOKVSYKGDPIVQ 300
QY 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVIKTCVOPTTKVRLLVKFPPELNY 360
D 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVIKTCVOPTTKVRLLVKFPPELNY 360
QY 361 QLKIKVCIDKDSGDVAALRGSRKRNILGTNTKVMNMEESNNNGSLSAEKKHLTLRQRCGN 420
D 361 QLKIKVCIDKDSGDVAALRGSRKRNILGTNTKVMNMEESNNNGSLSAEKKHLTLRQRCGN 420
QY 421 GGRANCDSALIVTEELHITFETEVYHOGKIDLETSHLPVVVISNICOMPMAASILMY 480
D 421 GGRANCDSALIVTEELHITFETEVYHOGKIDLETSHLPVVVISNICOMPMAASILMY 480
QY 481 NMLTNNPKNVNFFTPPIGTMDQVAEVLWSQFSSTTKRGLSIEQLTTIAEKILGCVNYS 540
D 481 NMLTNNPKNVNFFTPPIGTMDQVAEVLWSQFSSTTKRGLSIEQLTTIAEKILGCVNYS 540
QY 541 GCQITWAKFCCKENMAGKGFSEVWMLDNIIDLVKKYILALMNGCYIMGFSKREKALIST 600
D 541 GCQITWAKFCCKENMAGKGFSEVWMLDNIIDLVKKYILALMNGCYIMGFSKREKALIST 600
QY 601 KPPGFELLRFSESSKEGVTFTWVEKDISGKTQIOSVEPYTKOQLNNNSFAEIIINGYKIM 660
D 601 KPPGFELLRFSESSKEGVTFTWVEKDISGKTQIOSVEPYTKOQLNNNSFAEIIINGYKIM 660
QY 661 DATNILVSPLYLYPDITKEEAFGYKCRPESQEHPEADPGSAAPYLKTKFCVPTTCSN 720
D 661 DATNILVSPLYLYPDITKEEAFGYKCRPESQEHPEADPGSAAPYLKTKFCVPTTCSN 720
QY 721 TIDLPMSPRTDLSLMQFGNNGEAPESAGGOFESLTFPMDLTSSECATSPM 770
D 721 TIDLPMSPRTDLSLMQFGNNGEAPESAGGOFESLTFPMDLTSSECATSPM 770

RESULT 2
Q9BM54 PRELIMINARY: PRT: 769 AA.
AC Q9BM54:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DE SIGNAL PRODUCER AND ACTIVATOR OF TRANSCRIPTION 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC000627; AA00627.1; -
SQ SEQUENCE 769 AA: 87980 MW: A374A32AB9D28077 CRC64;
```

```
Query Match 99.1%; Score 3991.5; DB 4; Length 769;
Best Local Similarity 99.2%; Pred. No. 3e-262;
Matches 764; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 MAOMNLOOQLODTRVLEQJHOLYSDFPMELOKFLAPWIESODMAYAAKESHATLVPHNL 60
D 1 MAOMNLOOQLODTRVLEQJHOLYSDFPMELOKFLAPWIESODMAYAAKESHATLVPHNL 60
QY 61 LGEIDDOQYSRFLQESNNVLYOHNLKRIKQFLOSRYLEKPMELIARIVARCLMEESRLQTA 120
D 61 LGEIDDOQYSRFLQESNNVLYOHNLKRIKQFLOSRYLEKPMELIARIVARCLMEESRLQTA 120
QY 121 TAAOOGGAGNHPTAAVTEKQOOMEQHLQDVYKRVQDLEOKKVVENLODDFENYKTLK 180
D 121 TAAOOGGAGNHPTAAVTEKQOOMEQHLQDVYKRVQDLEOKKVVENLODDFENYKTLK 180
QY 181 SOGDMDLNGNNOSTYRQKMOOLEQMLTALDOMRSIYSELAGLLSAMEYVQKTLTDEEL 240
D 181 SOGDMDLNGNNOSTYRQKMOOLEQMLTALDOMRSIYSELAGLLSAMEYVQKTLTDEEL 240
QY 241 ADMKRREPETIACIGPPNICDLRENNITSLAESQLOTRQOIKKLEELQOKVSYKGDPIVQ 300
D 241 ADMKRREPETIACIGPPNICDLRENNITSLAESQLOTRQOIKKLEELQOKVSYKGDPIVQ 300
QY 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVIKTCVOPTTKVRLLVKFPPELNY 360
D 301 HRPMLEERIVELFRNLMSAFYVERQPCMPHDPRLVIKTCVOPTTKVRLLVKFPPELNY 360
QY 361 QLKIKVCIDKDSGDVAALRGSRKRNILGTNTKVMNMEESNNNGSLSAEKKHLTLRQRCGN 420
D 361 QLKIKVCIDKDSGDVAALRGSRKRNILGTNTKVMNMEESNNNGSLSAEKKHLTLRQRCGN 420
QY 421 GGRANCDSALIVTEELHITFETEVYHOGKIDLETSHLPVVVISNICOMPMAASILMY 480
D 421 GGRANCDSALIVTEELHITFETEVYHOGKIDLETSHLPVVVISNICOMPMAASILMY 480
QY 481 NMLTNNPKNVNFFTPPIGTMDQVAEVLWSQFSSTTKRGLSIEQLTTIAEKILGCVNYS 540
D 481 NMLTNNPKNVNFFTPPIGTMDQVAEVLWSQFSSTTKRGLSIEQLTTIAEKILGCVNYS 540
QY 541 GCQITWAKFCCKENMAGKGFSEVWMLDNIIDLVKKYILALMNGCYIMGFSKREKALIST 600
D 541 GCQITWAKFCCKENMAGKGFSEVWMLDNIIDLVKKYILALMNGCYIMGFSKREKALIST 600
QY 601 KPPGFELLRFSESSKEGVTFTWVEKDISGKTQIOSVEPYTKOQLNNNSFAEIIINGYKIM 660
D 601 KPPGFELLRFSESSKEGVTFTWVEKDISGKTQIOSVEPYTKOQLNNNSFAEIIINGYKIM 660
QY 661 DATNILVSPLYLYPDITKEEAFGYKCRPESQEHPEADPGSAAPYLKTKFCVPTTCSN 720
D 661 DATNILVSPLYLYPDITKEEAFGYKCRPESQEHPEADPGSAAPYLKTKFCVPTTCSN 720
QY 721 TIDLPMSPRTDLSLMQFGNNGEAPESAGGOFESLTFPMDLTSSECATSPM 770
D 721 TIDLPMSPRTDLSLMQFGNNGEAPESAGGOFESLTFPMDLTSSECATSPM 770

RESULT 3
Q9PYX8 PRELIMINARY: PRT: 769 AA.
AC Q9PYX8:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DE STAT 3.
GN STAT 3.
OS Xenopus laevis (African clawed frog)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
CX NCBI_TaxID=8335;
RN [1]
```


Oy	559	STKPGCTFLAIFSSSSNEGGCTFMWKEKDLSGRQLOISVDPYTKQOOLNNMSFPEIIMGX	658
Db	600	SPKPPTGCTFLAFSSSSKEGGCTFMWKEKDINGKTQIOISVEPYTKQOOLNMSFPEIIMGX	659
Oy	659	IMDATNLIVSPLVLYLPDIPKEEAFGYCYCPRESQEHEPA--DFGSAA-PYLTKTFICVT	714
Db	660	IMDATNLIVSPLVLYLPDIPKEEAFGYCYCREA--HDTFEPTDGTCYTGYLYLTKTFICVT	717
Oy	715	P-----TTCSNTIDL-PMSPRLTDSLMOFGUNNGCAEPDSAGGQFE	753
Db	718	PCPSVFMDPSEELGNGFCPTNGANTSDLEPMSPRLLDSL---HNEAAEAANP-GPLE	772
Oy	754	SLETFMDLTSECATSPM	770
Db	773	SLETLDELMSIDHA-SPM	788

```

RESULT      5
013133
AC      013133      PRELIMINARY;      PRT:      767 AA.
DT      01-JUL-1997 (TEMBLrel. 04, Created)
DT      01-JUL-1997 (TEMBLrel. 04, Last sequence update)
DT      01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE      STAT3.
CN      RSTSTAT3.
OS      Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC      Procladonhoperierygii; Salmoniformes; Salmonidae; Oncorhynchus.
CX      NCBI_TaxID=8022;
      ||
RN      SEQUENCE FROM N.A.
RP      Johnson M.C., Moutich D. V., Leong J. C.;
RL      Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR      EMBL: U60333; AAB60926.1;
DR      HSSP: P42227; 1BG1
DR      InterPro: IPR000980; SH2
DR      InterPro: IPR001217; STAT.
DR      Pfam: PF00017; SH2.1.
DR      Pfam: PF01017; STAT.1.
SU      SEQUENCE      767 AA; 87816 MW; FED97740C74C3798 CRC64;

```

Query Match	83.7%	Score 3373.5	DB 13	Length 767
Best local similarity	84.4%	Pred. No. 2.3e-220		
Matches	656	Conservative	53	Mismatches 51; Indels 17; Gaps 9.
QY	1	MAQWQLOLOLDRYYIKOLHQLYSDTFPMELRQFLAPMIESQWAAVAAKSESHATLVFNL	60	
DB	1	MAQWQLOLOLDRYYIKOLHQLYSDTFPMELRQFLAPMIESQWAAVAAKSESHATLVFNL	60	
QY	61	LGELIDQVSRFLQESNVLYOHNLRRIRKQFLQSGRYLEKPMELARIVARCLMEESRLLQTAA	120	
DB	61	LGELIDQVSRFLQESNVLYOHNLRRIRKQFLQSGRYLEKPMELARIVARCLMEESRLLQTAT	120	
QY	121	TAAGQCGGANNHTAAVVTTEKQOMLEOHLODVKRRQVDELQKKKVVENLDDHDFNYKTKL	180	
DB	121	TAAGQCGGANNHTAAVVTTEKQOMLEOHLODVKRRQVDELQKKKVVENLDDHDFNYKTKL	180	
QY	181	SQGDVODLNGNNOVSFROKMQOLEOMLTALDOKRRSIVSELGAGLSAMEVQVOTLIDE	238	
DB	180	SQGEISQDMKNSQAARTQKKSOLEOMHSALDOLRQVLTAMAGILSAMDVQKNTLDD	239	
QY	239	ELADMKRRPELACIGGPNICLDRLENNITSLAESOLOTRQOIKKLEELQOKVSYKGPDI	298	
DB	240	ELADMKRRQIACIGGPRICLDRLETWTSLGELIOLDRQOIKKLEELQOKVSYKGPDI	299	
QY	299	VQHRPRLERLYELFRNLKMSAFVVEROPCMHNDRLPVITGQVFTTKVALLKYPFL	358	
DB	300	IQHRPRLERLYELFRNLKMSAFVVEROPCMHNDRLPVITGQVFTTKVALLKYPFL	359	

OY	359	NYOLK IKVCJDKOSDVAALGSGKRENTLJGNTVMNNESSNNSJSAEKKHLTRPORC	418
Db	360	NYOLKIKVYDKESODVAAIKOSKRFNLLGNTKVMNNESSNNSJSAEKKHLTRPORC	419
OY	419	GNGGRANDASLIVTEELHLJTFETFEVYHOGKLIIDLETHSLPVPVVISNIGOMPNAMASIL	478
Db	420	GNGGRTNSDASLIVTEELHLJTFETFEVYHOGKLIIDLETHSLPVPVVISNIGOMPNAMASIL	479
OY	479	WYNMLTNNPKNVNPFYTKDPJGTMDOVAEVLWSMOFSSFTTKKGLSIEOL'TPLAEKLLGCVN	538
Db	480	WYNMLTNNPKNVNPFYTKDPJGVTMDOVAEVLWSMOFSSFTTKKGLTLEOLTTLAEKLLGCVN	539
OY	539	YSGCOITMAKPKCKEMMAKKGFSFWYMDLNIIDLVKYTLALANGNGYIMGFSIKERERAIL	598
Db	540	YSGCOITMAKPKCKEMMAKKGFSFWYMDLNIIDLVKYTLALANGNGYILGFIISKERERAIL	599
OY	599	STKPRGTFLTRPSSBSKSGGYFTWVEKDISGKTUISOVEPYTKOOLNNNSFAEIIIMGKY	658
Db	600	SKRPYTFELTRIESKKEGGYFTTWVEKDISGKTUISHVEPYTKOOLNNSJFAKIIIMGKY	659
OY	659	IMDATNIIIVSPILVLYLPDIPEKEARGKYCRESPOHPHPADPGSAA---PYLAKTKPICYVT	714
Db	660	IMDATNIIIVSPILVLYLPDKIPKEDAGNGYCRCREAA--PRAELGNGCITITOPYLAKTNFICVT	715
OY	715	PYTCSNTIDL-PMSPRTIDLSLMOFCONNEBGAEPASAGOFESULTPDMOLTSBCATSPM	770
Db	718	PNNSGNTSIDLPMASRTLNLSNI--HNNAENPN--GALJSLDILDMELISSIVA-SIIM	767

```

RESULT      6
O9QXKO
ID      O9QXKO      PRELIMINARY;      PRF;      749 AA.
AC      O9QXKO;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.
GN      STAT1.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OX      NCBI_TaxId=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Chen G., Newgard C.B.;
RT      *Expression of SPAT1 in INS-1 cells.*;
RL      Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR      EMBL: AF205604; AAF20200.1; -.
DR      HSSP: P42224; 1PF5.
DR      InterPro: IPR000980; SH2.
DR      InterPro: IPR001217; STAT.
DR      Pfam: PF00017; SH2; 1.
DR      Pfam: PF01017; STAT; 1.
DR      SMART: SM00252; SH2; 1.
DR      PROSITE: PS50001; SH2; 1.
SQ      SEQUENCE 749 AA: 87234 MW: D37C634215DEN355 CRC64;

```

Query Match	Similarity	Score	DB	Length
Best Local	51.8%	Pred. No. 7e-130		
Matches	396	Conservative	145	Mismatches 195; Indels 28; Gaps 11;
QY	1	MAOMNLOOLDRVYKLOLHQLYSDFFPMELROFLAPWIFSQMAYAAASKESHATLVFNL	60	
DB	1	MSOMFLOOLDSKFLFEOVHQLYDDSPMELIKOTLYLQWLEKQMEHAADVSATIRFIDL	60	
QY	61	IGEDIVQVSRFLQESNVLYLQHLRLRIKOFLOSRYLEKPMELARIARVACLMEESRLQTTAA	120	
DB	61	LSQLDQVSRFLQENFLFLQHLNIRKSRRLQDQFEDPQVQSMIIHNCLEKRLILEAQ	120	
QY	121	--TAADQGGANHPAAVATEKQOMLEQHLQVKKRVODILEQKMKVENVLQDDHDFENYKT	178	
DB	121	RFNQKQKQVIGIQN---TWALDKQKELDSKVSQVAVKQVQVMEIEHETFLDLDDDEVDFCKT	176	

Qy	179	LKQGDMDLNGNNSYTRQKMOLEOMLTALDQMRSSYSBELAGLSAMEYOKTLTDE	238
Db	177	--SONRESEANGVAKSDOKOEOQLLHFRFLMLDQKKREITHKIRELLNISIELYONTLIND	234
Qy	239	ELAWMKRPEELACIGCPNNICIDRLBEMWITSLSAQSLQTRQOIKKEELEQOKYSKGDPI	298
Db	235	ELVEEMKRRQGSACIGCPNNACLDQLOSMTIVANSIQOVRQOLKLEDEBQKFYEDPI	294
Qy	299	VQHRPMLFEERLEVELFRNIMKSAFVERQPCPHADRPVYKTVQTTKRYALLVKRPEL	358
Db	295	TKNQVLSDRFVFEVFOOLIOSSFVERQPCPHTRPORPLVKTQVETVKRLRLVYKLOEL	354
Qy	359	NYOLKIKVCIDKSDGDAVALGSRKRNILGNTKVMNMEESNCSLSAEFKHTLREORC	418
Db	355	NYNLKVKVSPKDVNENKNTVAGPKRPNILGHTKVMNMEESTNCSLSAEFRHLDQKQK -	413
Qy	419	GNGRANCDA SLVYTBELHLITFEETVEYHOGKLIKDELTHSLPYYVISNICOMPMANASIL	478
Db	414	-NAGKRTNEGPLVITHEELSHSFETOLQOPQLVIDLETTSLPYYVISNVSQLPQGMASIL	472
Qy	479	WYNMLTNPKNVNEFTKPPIDTMOQVAEVLISMOSSTPKRLSIEOLITTLAEKILGCVN	538
Db	473	WYNMLVTEPRNLSEFLNPPCMMWSQLSEVLISMOSSTPKRLNADQLSMLGEKILGPNAG	533
Qy	539	YSGCOITWAKFCENEMAGKGFSPVAVLDNITIDLVKKYILALNMGYIMGFSIKERERAIL	598
Db	533	PDGL-IPMTFRCKENINDKNPSFPMWIDTILELKKHLLCLMNDGCLMGFSIKERERAIL	591
Qy	599	SKTRPGFTLFFSESSKEGGVTFPWWKEDIS -GKTQIOSVEPYRKQOLNNNSPFIITMGY	657
Db	592	KDQOPGTFLLFSESSREGATFTFWERSQNGERPDHAYEPYTKKELSAVTFDITIRNY	651
Qy	658	KIMDATNIVSPLYLYPDIPKEEAFGK -YCRP -ESQHPHADPGSAAPYLKTKFCVT -	714
Db	652	KYMAENIPEBPLVLYLPNIDKHAFGKYRSRPEAEDEPMLDQPKRTGYITKELISVE	711
Qy	715	--PTTCSNTID-LPMSPRTDLSLMQFGNCGAPSGAGQGESL 755	
Db	712	VHPSRLQSTENLRLPMSPREPEMKI-----YGSFSDM 745	
RESULT	7		
Q9D323	09D323	PRELIMINARY: PRT: 749 AA.	
RC	STRAIN=C57BL/6J; TISSUE=COLON;		
RX	MEDLINE=21085660; PubMed=11217851;		
DT	01-JUN-2001 (TREMblrel. 17, Created)		
DT	01-JUN-2001 (TREMblrel. 17, Last sequence update)		
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.		
GN	START1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=COLON;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shihagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasudawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,		
RA	Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsi G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bolt C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofman M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Mordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wiltkeker C., Wilming L.,
RA Wyszynka-Borris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsunk S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK018544; BAB31265.1; -.
DR MGD; MGI:103063; Stc1l.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001217; SPAT.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; SPAT; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SO SEQUENCE 749 AA; 87280 MW; FT7D40AC9D9ED7C99 CMC64;

Query Match	Similarity	50.3%	Score 2026	DB 11	Length 749
Best Local	Similarity	51.7%	Prod. No. 4.6e129		
Matches	395	Conservative	144	Mismatches	197
				Indels	28
				Gaps	11
OY	1	MAOMNLOOJLDJRYVY,KOLHOLYSIPFPMEL,KOFLAPWYLSQWMAAYAAKSESHATVFNHIL	60		
OY	1	MSQWYFLIYVYKSLFLEQVHQLYDSDPFMEIKRYLQWMLKXQWMLAAATVYATATIRPHIL	60		
OY	61	LGEIDDOYSRLEQSSNLYQHNLRRIKQFLQSRYLEKPMELARIVAKLWEESSRLQVTA	120		
Db	61	LSOLDDOYSRSLFENNFFLIQHNIRKSKRMJQNFQEDDQVQSMILYNCLKEKKRDLLENKO	120		
OY	121	--TAAQGOAHHPAAVAAYTEKQMLEUQLQVRRKQVODLEQKMKVEMJODDFPNNKT	178		
Db	121	RRNQAOEGNTQ-----TYMLQKQKELDSKVRKVKQVQWYKIDEIKTELEJODEYDFCKCT	176		
OY	179	LKSQGDMODLNGNNSQVTKQKMOLEQMLTALDQMRRSIVSELACLSAMEYVQKTLTDE	238		
Db	177	--SQMRREGANVAASDQKOEQLLHKHETMLMDNKRKEILHRIRELLNSIELTQVTLIND	234		
OY	239	ELADKKRREIJCIGPPICDRLQENMTLSAESOLOTRQOIKRLEELQOKVSYKQDP1	298		
Db	235	ELYEKKRRQKQSCIGCPFRACIDQIQSMFTVAETLQOLRQOKLLELEQKFTTEPDP1	294		
OY	299	VOHRPMLERIVELERNLKMSAFVVERQPCMRPHDRPLVITQVQFTTKVRLVYKFPPEL	358		
Db	295	TKNKQVLSDFRFLFQOQLQSSFVEVERQPCMRPHDRPLVITQVQFTKVLTLVKLOEL	354		
OY	359	NVOLIKVCIDKSDQDVAALRGSKRFNLLGTPTKWMNNEESNGSLSAEPKHLTI,REORC	418		
Db	355	NINLVKVSFSDQDVEKNTVKGFRFNLGLTHTKWMNNEESNGSLAEFRHLQLEKQD-	413		
OY	419	GNGRANDASLIVELEHLITFEFEVYHQGLIDELEHSLPVTAVYISNICQPNMAASIL	478		
Db	414	-NAGRTEHGPILVTEELHLSLSEFTQLOCPGLVIDLETTSLPVTAVYISNVSQLPQSMASIL	472		
OY	479	WYNMLTNNPKVNFETTKPPIGTWDOVAEYLSQSFSSSTTKRGISIQULTTAAKLLGPGVN	538		
Db	473	WYNMLVTEPRNLSEFLPAPPCAWMSQSLSEVLSQSFSSVTKRGINAOLQSLMGLKELGLPGAG	532		
OY	539	YSGCOITAAKFCCKEMNAKGGFSEVWVLNIIIDLVKYYILALINEGYIMGTSIKERERALL	598		
Db	533	PDGL-IPMTRFCKEKNINQKNSFWMIWIDITIELIKHLLCLCMLNDQCMGIFISKERERALL	591		
OY	599	STKPPGTFLLRFSSESSKSGGVYFTTWVERQIS-GKTOIOSVEBYTKQOOLNNMSFAEITMGY	657		
Db	592	KDOQGTQFLLRFSSESSRGCAITFTFWVESQNGCEPDFAVEBYTKKELSAVFPD1IRNY	651		
OY	658	KIMDNTNLVSPLYVLDIPKEAEFGK-YQCP-ESQHPPEADPSAPYATKTRICAT-	714		
Db	652	KYMAAENIPENDKLKLYINIDKQHALGRTYYSRPKAPAPPEMELDDPKKRGYIKTELIVASE	711		
OY	715	--PTTCSMTID-LPMSPRTLDLSLMQFGNNGGAEPSSAGQGFESTL	755		
Db	712	VHPSRLQTTDNLNLPMSPEEPDEMSTI-----VGPEDSM	745		

RESULT 8
ID 099K94 PRELIMINARY; PRT: 712 AA.
AC 099K94:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC004808; AAH04808.1; -
SQ SEQUENCE 712 AA; 83106 MW; D62081709638CDA5 CRC64;

Query Match 49.8%; Score 2006; DB 11; Length 712;
Best Local Similarity 53.5%; Pred. No. 9,76-128;
Matches 385; Conservative 137; Mismatches 183; Indels 14; Gaps 8;

QY 1 MAONNOLOQDTRYLKOHLYSDTFPMEIROYLAIPWIESODMAYASKESHATLVFHNL 60
DB 1 MSOWPELOQLSKFLEOVHLYDSDFPMEIROYLAOWLEKODMEHAAYDVSATIRFHD 60
QY 61 LGEDDOYSRPLEOSNVLYOHNLRKIKOFLQSRYLEKPMETIARIYARCLWESRLLOTAA 120
DB 61 LSQDDOYSRPLESLENNFLQDHNIRKSRNLDNFDQEDPVQMSMIITNCKLEERKILENAO 120
QY 121 --TAOQGGQANHPPTAAVYTERKQOMLEOHLDVYRKRVODLEOKKAVENLODDPFPNKT 178
DB 121 RFNQAOEONIGN---TYMLDKOKELSDKVRNVQVQKICBOEIKTLELDEYDFKCT 176
QY 179 LKSGODMODLNNOSVYRQKMOOLEOMUTALDOMRSTVSELATLSAMEYVQCTLDE 238
DB 177 --SQNREBANGVAKSDOKOEDLLKMFMLDMKREKELIKRELNSIELTQWTLND 234
QY 239 ELADKKRPEIACIGPPNICLDRLNNMTSLAESOLQTRQOIKLELOOKVSYKGPDI 298
DB 235 ELVEKKRRQOSACIGPPNACIDQOSWFTTAEFTLOOIKLELEOKFTEYEPDI 294
QY 299 VOHRMLEERIVELPRNLMSAFYVEROPCMHPRDPLVKTGYOFTTVKLLVKFPEL 358
DB 295 TKNKOVLSDRFTFLFQOLQSSFVVEROPCMFTHPORPLVLTGVQFTVKLLVKLOEL 354
QY 359 NYOLKIKVCIDKSGDVAALRGSRKFNILGTNTKYNMESNNGSLSAEFKHLTREORC 418
DB 355 NYNLKVKVSFDKDVNEKMTVKGFRKFNILGTHTKYNMESNNGSLSAEFRLQKEDK 413
QY 419 GNGGANDASLIYTEELHLITFEVEYHOGKIDLETHSLPVVVITSNICOMPANASIL 478
DB 414 --NAGRTNEGPLIVTEELHLSFETQLCQPLVIDLETTSLPVVVITSNVSQISASIL 472
QY 479 WYNNLTNNPKVNFPTKPIGTMDQVAEVLISWQSSSTTKRGSLTLEQLTTLAEKLLGPGVN 538
DB 473 WYNNMLVTEPRNLSFPLNPGCAWMSQSLSEVLISWQSSSVTRGSLNQAOLSLGKLLGPNAG 532
QY 539 YSGCOITAAKCKENMAKGSFWWLDNIIDLVKYYILALMNEGYIMGFISKEERAIL 598
DB 533 PGDL-IPTTRCKENINKNFSFWPIDTLEIKKHLCLMNDCCIMGFISKEERAIL 591
QY 599 STKPGGTFLRPSSESKGCVFTWEKDIS-GKTOIOSVEPYTQOOLNNMSFARITNGY 657
DB 592 KDOQGTLLRPSSESRGCAITFTVERSQNGGEPDFAHVEYTKKELSAVTFPDIIRNY 651
QY 658 KIMDATNIVLSPVLYLPDIPEKAEGK-YCRP-ESQEHPEADPGSAAPYLTKFICVT 714
DB 652 KYMAENIPENPLKYLINIDKHAFGKYYSRPEAPERMELDDPKRTGYIKTELISVS 710

RESULT 9
ID 013131 PRELIMINARY; PRT: 754 AA.
AC 013131:
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIGNAL TRANSDUCER/ACTIVATOR OF TRANSCRIPTION STAT1.
GN RBTSTAT1-1.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnson M.C., Mourich D.V., Leong J.C.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U60331; AAB60924.1; -
DR HSSP: P42224; 1BF5.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 754 AA; 86872 MW; 9C88B812B0DC6777 CRC64;

Query Match 47.2%; Score 1901; DB 13; Length 754;
Best Local Similarity 51.4%; Pred. No. 1,4e-120;
Matches 380; Conservative 140; Mismatches 200; Indels 20; Gaps 12;

QY 1 MAONNOLOQDTRYLKOHLYSDTFPMEIROYLAIPWIESODMAYASKESHATLVFHNL 60
DB 1 MSOWPELOQLSKFLEOVHLYDSDFPMEIROYLAOWLEKODMEHAAYDVSATIRFHD 60
QY 61 LGEDDOYSRPLEOSNVLYOHNLRKIKOFLQSRYLEKPMETIARIYARCLWESRLLOTAA 120
DB 61 LAOLNQHRSRFALENNFLQDHNIRKSRNLDNFDQEDPVQMSMIITNCKLEERKILENAO 118
QY 121 TAOQGGQANHPPTAAVYTERKQOMLEOHLDVYRKRVODLEOKKAVENLODDPFPNKT 180
DB 121 LAOLNQHRSRFALENNFLQDHNIRKIKRNOHDHFOEDPVHMMIISKNLEKOKIL--AA 118
QY 179 LKSGODMODLNNOSVYRQKMOOLEOMUTALDOMRSTVSELATLSAMEYVQCTLDE 238
DB 177 NR-ETPMAGLTPKOLEHEKLLIAEMCKLKIKRGEVVGOLAEVLMMAEAVOSDLISEL 234
QY 241 ADMKKRPEIACIGPPNICLDRLNNMTSLAESOLQTRQOIKLELOOKVSYKGPDI 300
DB 235 ADMKKRPEIACIGPPNACIDQOLQWMTAAVASELDQVROQLKHQELQOKYITVDPDIKO 294
QY 301 HRPMLEERIVELPRNLMSAFYVEROPCMHPRDPLVKTGYOFTTVKLLVKFPEL 360
DB 295 OKGFLPRNATLEFRNLLEHLSLVEROPCMFTHPORPLVLTGVQFTVKLLVKLOEENY 354
QY 361 QLKIVCIDKSGDVAALRGSRKFNILGTNTKYNMESNNGSLSAEFKHLTREORC 420
DB 365 QLKIVCIDKSGDVAALRGSRKFNILGTNTKYNMESNNGSLSAEFKHLTREORC 409
QY 421 GGRANDASLIYTEELHLITFEVEYHOGKIDLETHSLPVVVITSNICOMPANASIL 480
DB 410 GNRTN-EGPLIVTEELHLSFETQLCQPLVIDLETTSLPVVVITSNVSQISASIL 468
QY 481 NMLTNNPKVNFPTKPIGTMDQVAEVLISWQSSSTTKRGSLTLEQLTTLAEKLLGPGV--N 538
DB 469 NMLTSEPKNLFELSPASWQSLSEVLISWQSSVYTKRGLNEQGLADKLGLGKAORN 528
QY 539 YSGCOITAAKCKENMAKGSFWWLDNIIDLVKYYILALMNEGYIMGFISKEERAIL 598
DB 529 PEGD-IPTTRCKENINKNFSFWPIDTLEIKKHLCLMNDCCIMGFISKEERAIL 586

QY	599	STKPGCTFLLPRESSKRGVTFWEKD	ISGKTQIOSVEPTKQOLNNMSPAEIIMGK	658
Db	567	TGKCGTFLLRSESSRGALFTTWEHDQINPVFAHVEPTIKELSAVS.LPDIKRTK		646
QY	659	IMDATNIIIVSPLYLYPDIPKEAFGKYC--RPESQEHPEADPGSADYBLKTKFCVPT		716
Db	647	VMAAENIPENPLRFLFXPDIPKDKSGKYYARASEASEPMDVGGSSSTGMYKTELISVSEV		706
QY	717	TCSNTID--LPMSPRTLDL		734
Db	707	HPSRLQDNMPMSPDVGGEL		726
RESULT	10			
ID	013132	PRELIMINARY:	PRT:	754 AA.
AC	013132:			
DT	01-JUL-1997	(TREMBLrel. 04, Created)		
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE	STAT1-2.			
GN	RBNSTAT1-2.			
OS	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;			
OC	Proclanchopterygii; Salmoniformes; Salmonidae; Oncorhynchus.			
OX	NCHI_TaxID=8022;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Johnson M.C., Mourich D.V., Leong J.C.;			
RL	Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: U60332; AAB60925.1; -.			
DR	HSSP: P42224; 1BF5.			
DR	InterPro: IPR000980; SH2.			
DR	InterPro: IPR001217; SPAT.			
DR	Pfam: PF00017; SH2; 1.			
DR	Pfam: PF01017; STAT; 1.			
DR	SMART: SM00252; SH2; 1.			
DR	PROSITE: PS50001; SH2; 1.			
QO	SEQUENCE 754 AA; 87147 MW; 340B9645EA040142 CRC64;			

[illegible]

Dd	355	QAKVAAALDGC---DVTEKKGFRRKFNILGNTNIVMMNEES--NOSLAAEFRRHLDLKEDQVA-	409
Oy	421	GGRANDASLIATVEELHLITFEYVYHOGLKTIDLETHSLPVYVINICOMPANAMASYLMY	480
Dd	410	GMRTN-EGFLIATEEJHCIFESELSNOSGLETKLETISLPYIVISNVSPISGMASIIIMY	468
Oy	481	NNLTNNPKRVNFETPPICGTMDVAEVLISMOSPSTFKRCLTSEOLTTLAEKLLGPV--N	538
Dd	469	NNLTSEPKNLAKFFLSPPAASHWGLSEVLISMOSPSSVTFRKLNEQLMLADKLIGCKAQRN	528
Oy	539	YSGCOITMAKPFCKENMAGKGFSFWMYLDNIIDLVKYIITALNEGVIIMCFISKERRALI.	598
Dd	529	PEGL-IPTMKFCK-SLSEKSFEPFMIIEMAILDLIRHLLSLWNDDCILLGVSEKEREKAL.	586
Oy	599	SRRPGCTFLRRSESSKEGGVFETWEKDISGKTOLQSYEPTTKOOLNMMSFAEIIMCYK	658
Dd	587	TOKCHOTFLNRSESSRDQAITYFTWBHLYXKPVFHAAPPTKRLSVALSLPDIIIRTYK	646
Oy	659	IMDATNIIVSPVLYLPDIPEKEAGKYC--RPSESHPEADPGSAAPYLKTRFCVTPPT	716
Dd	647	VMAERIENPENNPLRYLPDIPEKDSFGKYTYTRASASEPMDVGSSVGYMTELLISYSEV	706
Oy	717	TCSNTID--LPMSPRTDSL 734	
Dd	707	HPSRLQDNMPSPDPVFGEL 726	
RESULT	11		
ID	093598	PRELIMINARY:	PRT: 749 AA.
AC	093598:		
DT	01-NOV-1998	(TREMBLrel). 08. Created)	
DT	01-NOV-1998	(TREMBLrel). 08. Last sequence update)	
DE	01-JUN-2001	(TREMBLrel). 17. Last annotation update)	
GN	TRANSCRIPT: FACTOR.		
OS	Brachydontario (zebrafish) (zebra danio).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;		
OC	Cyprinifomes; Cyprinidae; Rasbora; Danio.		
RN	NBI_Taxid:7955;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RA	Oates A.C.,		
RL	Thesis (1998), University of Melbourne, Australia.		
DR	EMBL; AJ005642; CA06676.1; ..		
DR	HSSP; P422_1AF5.		
DR	InterPro: IPR00980; SH2.		
DR	InterPro: IPR01217; STAT.		
DR	Pfam: PF01217; SH2; 1.		
DR	Pfam: Pfam; SH2; 1.		
DR	SMART; Sm; SH2; 1.		
DR	PROSITE; PS50001; SH2; 1.		
SO	SEQUENCE 749 AA; 86856 MW; A84F66DEB2161137; NC664;		

Query Match	44.3%	Score 1786.5	DB 13	Length 749
Best Local Similarity	47.3%	Pred. No. 7.8e-113		
Matches	354	Conservative 159	Mismatches 203	Indels 33
				Gaps 15
QY	1	MAQNMNOLOOILPRVYIKOLHOLYSDFPMPMLROFLPAMWIESOOMAAVAAS--KESHATLVFN	58	
DB	1	MTQMLTFLQDLDRKFLBQDQLYDDNFPMALIKROYLSMIESHMDHVAANYENSLATVRN	60	
QY	59	NLLGETPDQYSFLEDSNVLVONHLNRRIKQFLQSRYLEKPMETARIVARCLWESRLQTT	118	
DB	61	DLTLQDLHDSHFVEKMEKFLNHNIRKFKRNQNLNFQDNPSVAMAMISNLRKEEKII--	118	
QY	119	AATAAQQSGQAHNPRAAVUTTEKQVLEQNLDOVRRKRUODLEQKMKVVEVLEQDDPPFNKT	178	
DB	119	-AIAQMSSEPRPNQDMLVEKNRREDAIKVATIKNNVQDAQETIKNLEQDQDEYDFKNKT	177	
QY	179	LKSGQDMQDLNG--NNQASTRQKMOQLEQMLTALLOMRRSIVSELAGLISAREVYQKTLT	236	

```
DB 178 YGSR---DELNGTISKEDLNREKL-MIHAFVYLNDKRMATISOKEVNLVPEVICELI 233
OY 237 DEELADMKRRREPEIACIGCPNPLICDLRENNWITSLAESQLOTRQOIKLEELQKVSYKD 296
DB 234 STELVEMKRRQOACIGCLTNVCLDQJONWFSTLGESLQSRQOLKLELEQKLSYSTD 293
OY 297 PIVQHRPMLEERIVELFRNLKSAFVEROPCMHDPRLVYKTCVOFTTKRLLVYKFP 356
DB 294 PTCOSKISLEERIMSNKLNVTNSLVEROPCMPTOLOPRLVLTGVLTKRLKRLVKIP 353
OY 357 ELVYQLIKICITIKDSGLVAALRGSRKFNILGTNTKYVNNNEESNGLSAEFKHLTLREQ 416
DB 354 EFNHTVAVKQOPKDI--VEKRGQFMENILGTSMKYVNNNEESN--CMAEFPRLQLEKER 409
OY 417 RCGNGRANCASLIYVEELHLITFEFEVYHQGLKIDLEFHSPLPVVYISNICOPNMAAS 476
DB 410 KV--TGMRTEGGLIVEELHJSTFEALCWSGLVINEFTTSLPIVVISNVSOPLSGMAS 467
OY 477 ILWYNNMLTNPNKNVNFETPKPIGTDQVAEVLISWQFSSTTKRGSLIBOLTTTLAEKILGP- 535
DB 468 IMYNNMLISEPKNLSTFVTPRPATWGOLOEVLSWQFSSTTKRGSLNPNQSLMLGKLLGPK 527
OY 536 GVNYSQGITWAKFCCKENMAGKGFSEVWMLDNIIDLVKYYILALMNEGYIMGFISKERER 595
DB 528 AASDFEAOIFMNNRCKSG-SEKNFTFWLMTLEAVLDLIKRLHLLSLMDGCIIMGFVTKDRTK 586
OY 596 AILSTRPGFTLLRFSSSEKGGVFTWVEKDISGTOIOSVEPYTKOOLNNSFAETIM 655
DB 587 SLNNKAPGFTLLRFSSSEKGGVFTWVEKDISGTOIOSVEPYTKOOLNNSFAETIM 646
OY 656 GYKIMDATNLLVSPVLYLPDIJPKKEAFGK-YCRPESQEHF---BADPGSAAPYLKTK 709
DB 647 NYRVYADNVPENLRELYLPQIPKDDAFKKYKTKPTDKQEMVDKRADEG---YISTP 702
OY 710 PICYT---PTTCSNTIDLPMSPTLDSL 734
DB 703 LITISEMKPOEYENL-MPMSPEVYGEL 729

RESULT 12
O9DDJ8 PRELIMINARY: PRT: 394 AA.
AC O9DDJ8:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE TRANSCRIPTION FACTOR STAT3 (FRAGMENT).
GN STAT3.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Miyoshi K., Cui Y., Riedinger G., Robinson P., Lehoczky J., Zon L.,
RA Oka T., Dewar K., Hemmighausen L.;
RT "Structure of the Mouse Stat 3/5 Locus Evolution from Drosophila to
RT Zebrafish to Mouse."
RL Genomics 0:0-0(2001).
DR EMBL: AF322857; AAC42495.1; -.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
FT NON_TER 1
SQ SEQUENCE 394 AA: 44126 MW: 28DEDEF716B50E70 CRC64;
```

Query Match 39.7%; Score 1597.5; DB 13; Length 394;

Best Local Similarity 82.6%; Pred. No. 2e-100; Matches 317; Conservative 13; Mismatches 21; Indels 33; Gaps 8;

```
OY 412 TLREORCGNGRANCASLIYVEELHLITFEFEVYHQGLKIDLEFHSPLPVVYISNICOMP 471
DB 1 TLREORCGNGRNCRTNSASLIYVEELHLITFEFEVYHQGLKIDLEFHSPLPVVYISNICOMP 60
OY 472 NAASTILMYNNMLTNPNKNVNFETPKPIGTDQVAEVLISWQFSSTTKRGSLIBOLTTTLAEK 531
DB 61 NAASTILMYNNMLTNPNKNVNFETPKPIGTDQVAEVLISWQFSSTTKRGSLIBOLTTTLAEK 120
OY 532 LLAGPGVNSQGITWAKFCCKENMAGKGFSEVWMLDNIIDLVKYYILALMNEGYIMGFISK 591
DB 121 LLAGPGVNSQGITWAKFCCKENMAGKGFSEVWMLDNIIDLVKYYILALMNEGYIMGFISK 180
OY 592 EERRALLSTPKPGFTLLRFSSSEKGGVFTWVEKDISGTOIOSVEPYTKOOLNNSFA 651
DB 181 EERRALLSTPKPGFTLLRFSSSEKGGVFTWVEKDISGTOIOSVEPYTKOOLNNSFA 240
OY 652 EILMGYKIMDATNLLVSPVLYLPDIJPKKEAFGK-YCRPESQEHF---DYGSAAPYLK 707
DB 241 EILMGYKIMDATNLLVSPVLYLPDIJPKKEAFGK-YCRPESQEHF---DYGSAAPYLK 298
OY 708 TKFLCYTP-----TTCSNTIDLPMSPTLDSL 746
DB 299 TKFLCYTPCPSVFWDFPDSLLGNGFPGTNSGNTSLDFPMSPTLDSL---HNEAFA 354
OY 747 SAGGPFESTLFPMDLTSECATSPM 770
DB 355 NP-GPLESLTLDMLSSDHA-SPM 376

RESULT 13
O9QXJ2 PRELIMINARY: PRT: 922 AA.
AC O9QXJ2:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.
GN STAT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Farrar J.D., Smith J.D., Murphy T.L., Murphy K.M.;
RC STRAIN=B10.A;
RT "Sequence Divergence in the C-Termini of Murine and Human Stat2
RT Controls Species-Specific Stat4 Activation and Th1 Development Induced
RT by Type I Interferons."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF206162; AAF17096.1; -.
DR HSSP: P42224; 1BF5.
DR MGD: MGI:103039; Stat2.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 922 AA: 105365 MW: 61E73D3276D5354A CRC64;
```

Query Match 31.9%; Score 1287; DB 11; Length 922;
Best Local Similarity 39.3%; Pred. No. 7.3e-79; Matches 294; Conservative 140; Mismatches 259; Indels 56; Gaps 20;

>C Eukaryota: Metazoa Chordata: Craniata: Vertebrata: Euteleostomi:
>OC Actinopterygii: Neopterygii: Teleostei: Euteleostei:
>OC Protacanthopterygii: Salmoniformes: Salmonidae: Oncorhynchus.
>OX NCBI_TaxId=8022;
>RN [1]
>RP SEQUENCE FROM N.A.
>RA Johnson M.C., Mourich D.V., Leong J.C.;
>RT "Rainbow Trout STAT Genes: Evolutionary Duplication Leads to STAT 5A
>RT and the Acquisition of Mammalian Specific Function."
>RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
>DR EMBL: AF147499; AAG14946.1; -
>DR InterPro: IPR000980; SH2; -
>DR InterPro: IPR001217; STAT.
>DR Pfam: PF00017; SH2; 1.
>DR Pfam: PF01017; STAT; 1.
>DR SMART: SM00252; SH2; 1.
>DR PROSITE: PS50001; SH2; 1.
>SQ SEQUENCE 784 AA: 89780 MW: 1C16DCFC7C83AF7C CRC64;

Query Match 21.7%; Score 873; DB 13; Length 784;
Best Local Similarity 30.4%; Pred. No. 6,6e-51;
Matches 248; Conservative 151; Mismatches 314; Indels 104; Gaps 29;

OY 1 MAOMNLOOLDPRYLKQLHQLYSDFRPMELROFLAPMIESODMAYV---ASKESHATLY 56
DB 1 MAVVIOAOQLQDGLAHQMSLYGQHFPIEVRYHLSQMLDQMLDVIDLENPOEEFKAKRL 60
OY 57 PNLGELIDQOYSRFLQESNVLYQHNLRRIKQFLOSRYLEKPMETARIVARCLMESRLL 116
DB 61 LDLSLQELONKKAHEGDEGFLKIKIGHYASQLKSTYDRCPLEIVRCIKHILYTEQRLV 120
OY 117 QTAATAAQ--CGOANHPTAAVTEKQMLQHLQDYRKRVQDLQKMKYVENLQDDPDFN 175
DB 121 REASNSSPVGGKMD--MSQKYQIQAFAEELRLTODTENDLRKIQHNQEXFYIIQ 175
OY 176 Y-KTKSKQGMODL-----NGNNSVTRKMQOLEQMLT-----ALQMRRSIVSELG 223
DB 176 YQESLRIGQNLSTLTLPAPDQRLREPTLSKRAIYEAWLITREANTLOKRYDLAKHOK 235
OY 224 LLSAMEYVOKLTLEELADMKRPEIACIGRPNICDLRLNMTSLAESQLOTROOIKK 283
DB 236 SLQLRKQOTIILDELIMHKRROQLAGNGGPEGGLDIQSCEKLAETIMQNRQOIRR 295
OY 284 LEELQOKVSKGDPYVQHRPMLERIVELFRNLMSAFVERQPCMHDPRIYIKGV 343
DB 296 AEHLRQQLPIG-PIEELINDLNTITTDIISALVSTFIIEKOP-----PQVLKTOT 346
OY 344 QFTKVRLLVKEPELNYOL---KIKVICIDKSGDVAAALRSRKN-----ILGTNTKVMN 395
DB 347 KEAATVRLVG--GKLNVHNPQVKATIISEQAKALLKNENMRNDSGEILNNC--VME 404
OY 396 MEESNNGSLSAEFKHLTLREQCGNGRANCASLIYTELHLITFEYVHOG--LKID 453
DB 405 YHQV--TGLTISANFRNMSLK--RIKRSDRGAES--VTEKEFTILFESQFSVQGNELVQ 458
OY 454 LEHSLPVVVISNICMPNMAASILVYNMLTNPKVNFETKPRIGTMDQVAEVLMSQFS 513
DB 459 VKTLSLPVVVYIHGSDNNATATVMDNAFA--EPGRVPTIV--PDKVLMPOLCEALNMKTK 516
OY 514 S--TTKRGISIEQLTLAEKLLGPV---NYSQOITWAKFCKENMAGKGFSPVWMLDN 567
DB 517 AEVQSNRGLSEENLYVLAQKAFSSSINPDYRGMTMTWSQFNRESLPGRNFTFMQWFDG 576
OY 568 IIDLVKKYILALMNEGYINGFISKEREKAILSTKPRGFTLLRFSSESKGCVTTFWEKD 627
DB 577 VVELTKRKHLPKPMNDALIGFVNRQAOADMLSKPNGTFLRLRFS--SEIGGITIAVVAEN 635
OY 628 IS--GKTOISQVEPYTKOOLNMNSFAEIIIMGYKIMDATNILVSLVYLVPDIPKEEAPGK 685
DB 636 PKKPGERWVWNLMPYTKKFSIRSLAD-----RISDLNHL-----FLYPRDKRDEVFESK 685
OY 686 YCRPESQEHPEADPGSAA--PYLKTKEICVTPTCSENTIDLPMSPTIDSLMQFGNNGEG 743

DB 686 YTP-----PLSKAVDGIVKPOIKOVVPEFTTTNPDPAANPTVYU-----HG 727
OY 744 AEPSAG-----GQFESLT-----FDMULTSECA 766
DB 728 ASPVSHPRPNYGIYQPMDSIILDADGDFOLDPTMDVA 764

Search completed: March 19, 2002, 15:25:44
Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 19, 2002, 15:19:38 ; Search time 23.78 Seconds

(Without alignments)
1187.213 Million cell updates/sec

Title: US-08-212-185-12

Perfect score: 4029

Sequence: 1 MAQWNLQQLDRLRYLKQLHQ.....QFESLTFMDLTSECATSPM 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4010	99.5	770	1	STRA3_MOUSE
2	4007	99.5	770	1	STRA3_MOUSE
3	3977	98.7	770	1	STRA3_MOUSE
4	2017	50.1	750	1	STRA1_HUMAN
5	1984	49.2	749	1	STRA1_MOUSE
6	1798	44.6	748	1	STRA4_MOUSE
7	1780.5	44.2	749	1	STRA4_MOUSE
8	1293	32.1	851	1	STRA2_HUMAN
9	1271.5	31.6	923	1	STRA2_MOUSE
10	1264.5	31.4	864	1	STRA2_MOUSE
11	867	21.5	799	1	STRA2_MOUSE
12	865.5	21.5	786	1	STRA2_MOUSE
13	865.5	21.5	786	1	STRA2_MOUSE
14	862.5	21.4	786	1	STRA2_MOUSE
15	862.5	21.4	786	1	STRA2_MOUSE
16	859	21.3	787	1	STRA2_MOUSE
17	858.5	21.3	794	1	STRA2_MOUSE
18	858	21.3	793	1	STRA2_MOUSE
19	857	21.3	793	1	STRA2_MOUSE
20	855	21.2	787	1	STRA2_MOUSE
21	791.5	19.6	794	1	STRA2_MOUSE
22	582	13.9	837	1	STRA6_MOUSE
23	558.5	13.9	847	1	STRA6_MOUSE
24	481	11.9	761	1	STAT_DROME
25	159	3.9	1085	1	RA26_YEAST
26	148	3.7	2663	1	CENE_HUMAN
27	145	3.6	1427	1	REST_HUMAN
28	145	3.6	1433	1	REST_HUMAN
29	145	3.6	1790	1	UISO1_YEAST
30	144.5	3.6	1959	1	MYH9_CHICK
31	143	3.5	1087	1	AKR9_RABIT
32	141	3.5	3911	1	AKR9_HUMAN
33	140.5	3.5	1312	1	RA50_YEAST

34	140.5	3.5	1526	1	MS2_SCHPO	O9u516 schizosacch
35	140.5	3.5	2230	1	G04_HUMAN	O13439 homo sapien
36	140	3.5	1940	1	MYH3_CHICK	P02565 gallus galli
37	140	3.5	3685	1	DMD_HUMAN	P11532 homo sapien
38	138.5	3.4	1939	1	MYH1_HUMAN	P12882 homo sapien
39	138	3.4	2411	1	MYSA_DROME	P05661 drosophila
40	137.5	3.4	1818	1	HMW2_MYCPN	P75471 mycoplasma
41	137	3.4	1978	1	MYH8_CHICK	P10587 gallus galli
42	136	3.4	857	1	MEDE_GITALA	O08014 garteria lam
43	135.5	3.4	1505	1	CUT1_HUMAN	P39880 homo sapien
44	135.5	3.4	1937	1	MYH8_HUMAN	P13535 homo sapien
45	135.5	3.4	2349	1	TPR_HUMAN	P12270 homo sapien

ALIGNMENTS

RESULT 1	STRA3_MOUSE	STANDARD:	PRT:	770 AA.
ID	STRA3_MOUSE			
AC	P42227:			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3 (ACUTE-PHASE RESPONSE FACTOR).			
DE	START3 OR ABRF.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 154-158; 181-185 AND 632-640.			
RC	STRAIN-BALB/C; TISSUE=Liver;			
RX	MEDLINE=94208062; PubMed=7512451;			
RA	Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,			
RA	Yoshida K., Sudo T., Naruto M., Kishimoto T.;			
RT	"Molecular cloning of ABRF, a novel IFN-stimulated gene factor 3 p91-			
RT	related transcription factor involved in the gp130-mediated signaling			
RT	pathway.";			
RL	Cell 77:63-71(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Thymus;			
RX	MEDLINE=94188718; PubMed=8140422;			
RA	Zhong Z., Wen Z., Darnell J.E. Jr.;			
RT	"Stat3: a Stat family member activated by tyrosine phosphorylation in			
RT	response to epidermal growth factor and interleukin-6.";			
RL	Science 264:95-98(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=95014185; PubMed=7523373;			
RA	Raz R., Durbin J.E., Levy D.E.;			
RT	"Acute phase response factor and additional members of the			
RT	interferon-stimulated gene factor 3 family integrate diverse signals			
RT	from cytokines, interferons, and growth factors.";			
RL	J. Biol. Chem. 269:24391-24395(1994).			
RN	[4]			
RP	SEQUENCE FROM N.A. (STAT3B).			
RC	STRAIN-BALB/C, AND B6; TISSUE=Liver;			
RX	MEDLINE=96016116; PubMed=7568080;			
RA	Schaefer T.S., Sanders L.K., Nathans D.;			
RT	"Cooperative transcriptional activity of Jun and Stat3 beta, a short			
RT	form of Stat3.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).			
RN	[5]			
RP	PHOSPHORYLATION OF SER-727, AND MUTAGENESIS.			
RX	MEDLINE=95354205; PubMed=7543024;			
RA	Wen Z., Zhong Z., Darnell J.E. Jr.;			
RT	"Maximal activation of transcription by Stat1 and Stat3 requires both			
RT	tyrosine and serine phosphorylation.";			
RL	Cell 82:241-250(1995).			

RN	[6]
RP	X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
RX	MEDLINE:98334373; Pubmed=9671298;
RT	Becker S., Groner B., Mueller C.W.:
KA	"three-dimensional structure of the Stat3beta homodimer bound to DNA.";
RL	Nature 394:145-151(1998).
CC	-I- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6
CC	(IL-6)-RESPONSIVE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS
CC	ACUTE-PHASE PROTEIN GENES. STAT3B INTERACTS WITH THE N-TERMINAL
CC	PART OF C-JUN TO ACTIVATE SUCH PROMOTERS IN A COOPERATIVE WAY .
CC	-I- PATHWAY: INVOLVED IN THE G1/SO-MEDIATED SIGNALING PATHWAY .
CC	-I- SUBUNIT: FORMS A HOMODIMER OR A HETERO-DIMER WITH A RELATED FAMILY
CC	MEMBER (AT LEAST STAT1) (BY SIMILARITY)
CC	-I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC	RESPONSE TO PHOSPHORYLATION.
CC	-I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; STAT3A (SHOWN HERE) AND STAT3B;
CC	ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	-I- TISSUE SPECIFICITY: STAT3A IS SEEN IN THE LIVER, SPLEEN, AND
CC	KIDNEY. STAT3B IS ALSO DETECTED IN THE LIVER, ALTHOUGH IN A MUCH
CC	LESS ABUNDANT MANNER.
CC	-I- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNTF,
CC	LIF, CSF-1, EGF, PGDF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION
CC	IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3
CC	HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC	-I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC	-I- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: L29278; AAA37254.1; -
DR	EMBL: U06922; AAA19452.1; -
DR	EMBL: U08378; AAA56668.1; -
DR	EMBL: U30709; AAC52612.1; -
DR	PDB: 1BG1; 1J-JAN-99.
DR	TRANSFAC: T01574; -
DR	MCD: MG1:103038; stat3c.
DR	InterPro: IPR000980; SH2.
DR	InterPro: IPR001217; STAT.
DR	Pfam: PF00017; SH2_1.
DR	Pfam: PF01017; STAT_1.
DR	SMART: SM00252; SH2_1.
DR	PROSITE: PS00011; SH2; 1.
KW	Transcription regulation; DNA-binding; Nuclear protein; SH2 domain;
KW	Phosphorylation; Alternative splicing; 3D-structure.
FT	DOMAIN 580 670
FT	MOD_RES 705 705
FT	PHOSPHORYLATION (BY JAKs) (BY
FT	SIMILARITY).
FT	MOD_RES 727 727
FT	VARSPLIC 716 770
FT	PHOSPHORYLATION.
FT	TTCNTIDLPSPTSLDSLMFGNGNGSGAESAGQESLSL
FT	FDMDLTSFCATSPM -> FLDAYMK (IN ISOFORM
FT	STAT3B).
FT	MISSING (IN SOME CLONES).
FT	S->A: DECREASED TRANSCRIPTIONAL
FT	ACTIVATION.
FT	E -> K (IN REF. 2).
FT	S -> T (IN REF. 2 AND 4).
FT	M -> I (IN REF. 1).
FT	CONFLICT 394 394
FT	CONFLICT 770 AA: 88053 MW: 6C00626711C8012D CRC64:
SO	SEQUENCE

Query Match	99.5%	Score 4010;	DB 1;	Length 770;
Best Local Similarity	99.5%;	Pred. No. 8.4e-244;		
Matches 766;	Conservative 3;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	MAQWQLOQLDPRYLKQHLXSDFFPHELRLQFLAPWIESQWYAAAKESATLVFNNL	60	

D	b	1	MAONHMOLOQDLTRRYLEQHOLYSDBSFPEHLEQFLAPMIESODMAYAAKESHATLVFINL	60
Q	y	61	LGEIDDOYSRFLQESNVL.YOHNLRR1KOPLOSRYLEKPMELIAIVAK1LMEESKRLQTA	120
D	b	61	LGEIDDOYSRFLQESNVL.YOHNLRR1KOPLOSRYLEKPMELIAIVARCLMEESKRLQTA	120
Q	y	121	TAACGGGGAHNPTAAVAVTEKOOMLQEHLODYKRRVQULQEQMKVVENQJWFDNRYTKL	180
D	b	121	TAACGGGGAHNPTAAVAVTEKOOMLQEHLODYKRRVQULQEQMKVVENQJWFDNRYTKL	180
Q	y	181	SGODQODLNGNNOVSTYRKMOMLEQMLTALQMRKVSIVSEIAGLILSAMEYQKTLTJPEL	240
D	b	181	SGODQODLNGNNOVSTYRKMOMLEQMLTALQMRKVSIVSEIAGLILSAMEYQKTLTJPEL	240
Q	y	241	ADMKRPELACIGCPNPLCLDKLEMM1TSLAESQJQTLMOQIKKLEELQOKKSYGDP1VQ	300
D	b	241	ADMKRPELACIGCPNPLCLDKLEMM1TSLAESQJQTLMOQIKKLEELQOKKSYGDP1VQ	300
Q	y	301	HRPMLSEERIVELFRILMKSAFVVERQPCMPHMDRPLV1KGVQPTTKVRILVPELNY	360
D	b	301	HRPMLSEERIVELFRILMKSAFVVERQPCMPHMDRPLV1KGVQPTTKVRILVPELNY	360
Q	y	361	OLK1KVCIDKOSGDVAALRGSRKFNI1LCTNTRKVNMEESNNGS1SAEPKHLTLNEQRCG	420
D	b	361	OLK1KVCIDKOSGDVAALRGSRKFNI1LCTNTRKVNMEESNNGS1SAEPKHLTLNEQRCG	420
Q	y	421	GGRANCDSLIVTEELH1ITFETEVYHOGK1DLETHSLPVVVISN1COMPMANAS1LMY	480
D	b	421	GGRANCDSLIVTEELH1ITFETEVYHOGK1DLETHSLPVVVISN1COMPMANAS1LMY	480
Q	y	481	NMLTNPNKRVNEFTKPP1G1TMDQVAEVL5WQFSSTTRKGLS1EQLTTLAEKLLGPGVNS	540
D	b	481	NMLTNPNKRVNEFTKPP1G1TMDQVAEVL5WQFSSTTRKGLS1EQLTTLAEKLLGPGVNS	540
Q	y	541	GCOTWAFCKENMAGKGF5FWMLDN1IDLKYY1IALNMGY1MGTSERERALL1ST	600
D	b	541	GCOTWAFCKENMAGKGF5FWMLDN1IDLKYY1IALNMGY1MGTSERERALL1ST	600
Q	y	601	KPPGFTLLRFSSSKSKEGVTFTWWEKD1SGKTO1QSVPEPYTKOOLNNNSFAE1IMGYK1M	660
D	b	601	KPPGFTLLRFSSSKSKEGVTFTWWEKD1SGKTO1QSVPEPYTKOOLNNNSFAE1IMGYK1M	660
Q	y	661	DATNLV5PLVLYLPD1PDKPEAFKCYKCPSEOB1PEADPGSAAPYLKTKF1CVPTPTCSN	720
D	b	661	DATNLV5PLVLYLPD1PDKPEAFKCYKCPSEOB1PEADPGSAAPYLKTKF1CVPTPTCSN	720
Q	y	721	T1D1PMSRPTLDS1LQPGNCKGAPRSAGGQFES1TFTPMDLTSSECATSPM	770
D	b	721	T1D1PMSRPTLDS1LQPGNCKGAPRSAGGQFES1TFTPMDLTSSECATSPM	770
RESULT 2				
STAT3_RAT				
ID	STAT3_RAT	STANDARD:	PRT:	770 AA.
AC	P52631:			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3.			
GN	STAT3.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxId=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver.			
RC	MEDLINE=96102059: PubMed=8530402:			
RA	Rippberger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,			
RA	Fey G.H.;			
RT	"Transcription factors Stat3 and Stat5b are present in rat liver			
RT	nuclei late in an acute phase response and bind interleukin-6			
RT	response elements ";			

```

RL J. Biol. Chem. 270:29998-30006(1995).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6
CC (IL-6)-RESPONSIVE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS
CC ACUTE-PHASE PROTEIN GENES.
CC -1- PATHWAY: INVOLVED IN THE Gp130-MEDIATED SIGNALING PATHWAY.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (AT LEAST STAT1) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNTF,
CC LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION
CC IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3
CC HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X91810: CAA62920.1;
DR HSSP: P42224: 1BFS5
DR InterPro: IPR000980: SH2
DR InterPro: IPR001217: STAT.
DR Pfam: PF00017: SH2; 1.
DR SMART: SM00252: SH2; 1.
DR PROSITE: PS50001: SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT MOD_RES 580 670 SH2.
FT DOMAIN 580 705 PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).
FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
FT FT SEQUENCE 770 AA; 86039 MW; D74A0C76954754ED CRC64;
SQ

```

Query Match 99.5%; Score 4007; DB 1; Length 770;
 Best Local Similarity 99.4%; Pred. No. 1.3e-243;
 Matches 765; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 MAQMNOLQQLDTRRYLQHLQYSDTFPEMLRQFLAPNIESQDMAYAAKSKESHATLVFHNL 60
DB 1 MAQMNOLQQLDTRRYLQHLQYSDTFPEMLRQFLAPNIESQDMAYAAKSKESHATLVFHNL 60
OY 1 LGELIDQYSRFLQESNVLYOHNLRRIKQFLQSRLEKPMELARIVARCLMEESRLQATA 120
DB 1 LGELIDQYSRFLQESNVLYOHNLRRIKQFLQSRLEKPMELARIVARCLMEESRLQATA 120
OY 121 TAAQGGQGANHPPTAAVYTEKQMLEOHLDVRRKRVODLEQKMKVENVLQDDDFENYTKL 180
DB 121 TAAQGGQGANHPPTAAVYTEKQMLEOHLDVRRKRVODLEQKMKVENVLQDDDFENYTKL 180
OY 121 TAAQGGQGANHPPTAAVYTEKQMLEOHLDVRRKRVODLEQKMKVENVLQDDDFENYTKL 180
DB 121 TAAQGGQGANHPPTAAVYTEKQMLEOHLDVRRKRVODLEQKMKVENVLQDDDFENYTKL 180
OY 181 SQGDMODLNGNNSVTRQKQMLEOMLTALDQMRISVSELAGILSMEYVQKTLTDEL 240
DB 181 SQGDMODLNGNNSVTRQKQMLEOMLTALDQMRISVSELAGILSMEYVQKTLTDEL 240
OY 181 SQGDMODLNGNNSVTRQKQMLEOMLTALDQMRISVSELAGILSMEYVQKTLTDEL 240
DB 181 SQGDMODLNGNNSVTRQKQMLEOMLTALDQMRISVSELAGILSMEYVQKTLTDEL 240
OY 241 ADKRRPETAICIGPPNICIDRLNNITSLAESOLOTRQOIKKLELQOKVSKGDPVYQ 300
DB 241 ADKRRPETAICIGPPNICIDRLNNITSLAESOLOTRQOIKKLELQOKVSKGDPVYQ 300
OY 241 ADKRRPETAICIGPPNICIDRLNNITSLAESOLOTRQOIKKLELQOKVSKGDPVYQ 300
DB 241 ADKRRPETAICIGPPNICIDRLNNITSLAESOLOTRQOIKKLELQOKVSKGDPVYQ 300
OY 301 HRPMLEBERIVLEFRNLKMSAFVVEROPCPMPHDPRLVITGVQFTTKVRLVKFPELNY 360
DB 301 HRPMLEBERIVLEFRNLKMSAFVVEROPCPMPHDPRLVITGVQFTTKVRLVKFPELNY 360
OY 301 HRPMLEBERIVLEFRNLKMSAFVVEROPCPMPHDPRLVITGVQFTTKVRLVKFPELNY 360
DB 301 HRPMLEBERIVLEFRNLKMSAFVVEROPCPMPHDPRLVITGVQFTTKVRLVKFPELNY 360
OY 361 QLKIKVICDKSDVAALKSGSKFNLGLTNTKYMMNEESNGSLSAEFKHLTLREORCGN 420
DB 361 QLKIKVICDKSDVAALKSGSKFNLGLTNTKYMMNEESNGSLSAEFKHLTLREORCGN 420

```

```

OY 421 CGRANCASLIVTEBLHITFETVEYHUGLKIDLETHSLPVVVISNICOMPAMASILMY 480
DB 421 CGRANCASLIVTEBLHITFETVEYHUGLKIDLETHSLPVVVISNICOMPAMASILMY 480
OY 481 NMUTNPNKVNPFKPPIGTDQVAEVLSSKQPSSTTKGLSIEQLTLAEKLLGCVNYS 540
DB 481 NMUTNPNKVNPFKPPIGTDQVAEVLSSKQPSSTTKGLSIEQLTLAEKLLGCVNYS 540
OY 541 GCOITMAFCEENNAKGFSEFWMLDNIIDLKRYITLALMEGYTMGFIISKEREALIST 600
DB 541 GCOITMAFCEENNAKGFSEFWMLDNIIDLKRYITLALMEGYTMGFIISKEREALIST 600
OY 601 KPGFTFLRFSESSKEGCVTFETWEKDISGRTVQISVEPYTKOOLNNNSFAELIMGYKIM 660
DB 601 KPGFTFLRFSESSKEGCVTFETWEKDISGRTVQISVEPYTKOOLNNNSFAELIMGYKIM 660
OY 661 DATNIVLSPLYLVDIPKEAFKCYCPSEQDEHPGSAAPYLKTKFICVPTTCSN 720
DB 661 DATNIVLSPLYLVDIPKEAFKCYCPSEQDEHPGSAAPYLKTKFICVPTTCSN 720
OY 721 TIDLPMSPTLDSLMQFGNNGEAGPSAGOFESITFPMDLTSECATSPM 770
DB 721 TIDLPMSPTLDSLMQFGNNGEAGPSAGOFESITFPMDLTSECATSPM 770

```

RESULT 3
 STAT3 HUMAN STANDARD; PRT; 770 AA.
 ID STAT3_HUMAN P40763;
 AC P40763;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 3 (ACUTE-PHASE
 DE RESPONSE FACTOR).
 GN STAT3 OR ATRF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_Taxid 9606;
 RN [1]
 RP SEQUENCE DATA: N. A.
 RC TISSUE: placenta.
 RX MEDLINE=94208062; Pubmed=7512451;
 RA Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
 RT Yoshida K., Suho T., Naruto M., Kishimoto T.;
 RT "Molecular cloning of ATRF, a novel IFN-gamma-induced gene factor 3 p91-
 RT related transcription factor involved in the IFN-gamma mediated signaling
 RT pathway.";
 RL Cell 77:61-69 (1994).
 RN [2]
 RP PHOSPHORYLATION: ON SERINE.
 RX MEDLINE=94208062; Pubmed=7701321;
 RA Zhang X., Zhang J., Li H.-C., Schindler C., Chen-Kiang S.;
 RT "Requirement of serine phosphorylation for formation of STAT-promoter
 RT complexes.";
 RL Science 267:11990-11994 (1995).
 CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE INTERLEUKIN-6
 CC (IL-6)-RESPONSIVE ELEMENTS IDENTIFIED IN THE PROMOTERS OF VARIOUS
 CC ACUTE-PHASE PROTEIN GENES.
 CC -1- PATHWAY: INVOLVED IN THE Gp130-MEDIATED SIGNALING PATHWAY.
 CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
 CC MEMBER (AT LEAST STAT1).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION.
 CC -1- TISSUE SPECIFICITY: HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL
 CC MUSCLE, KIDNEY, AND PANCREAS.
 CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-6, IL-11, CNTF,
 CC LIF, CSF-1, EGF, PDGF, IFN-ALPHA AND OSM. SERINE PHOSPHORYLATION
 CC IS IMPORTANT FOR THE FORMATION OF STABLE DNA-BINDING STAT3
 CC HOMODIMERS AND MAXIMAL TRANSCRIPTIONAL ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L29277; AAA58374.1; -
DR HSSP: P42224; 1BF5.
DR TRASNPAAC: T01493; -
DR MIM: 102582; -
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SMO0252; SH2; 1.
DR PROSITE: PSS0001; SH2; 1.
KM Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 580 670 SH2.
FT MOD_RES 705 705 PHOSPHORYLATION (BY JAKS) (BY
FT MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 770 AA: 88052 MW: A3DCEB815B3B5360 CRC64:

Query Match 98.7%; Score 3977; DB 1; Length 770;
Best Local Similarity 98.6%; Pred. No. 9,8e-242;
Matches 759; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 MAQNNQLOQDTRYLKQHOHLYSDFFPMELOFLAPWIESQDMAVAASKESATLVFNNL 60
DB 1 MAQNNQLOQDTRYLKQHOHLYSDFFPMELOFLAPWIESQDMAVAASKESATLVFNNL 60
QY 61 LGEDIDQYSRFLQESNVLYOHNLRIKQFLOSRYLEKPMETARIYARCLMEBSRLQTA 120
DB 61 LGEDIDQYSRFLQESNVLYOHNLRIKQFLOSRYLEKPMETARIYARCLMEBSRLQTA 120
QY 121 TAAQGGAGNHPPTAAVTEKQMLEQHLQDYRKRVQDLEQKMKVVENLQDDFDFNYKTLK 180
DB 121 TAAQGGAGNHPPTAAVTEKQMLEQHLQDYRKRVQDLEQKMKVVENLQDDFDFNYKTLK 180
QY 181 SOGMODLNGNNOSTYRQKMOLEQMLTALDOMRSIVSELAGLISAMEYOKTLTDEEL 240
DB 181 SOGMODLNGNNOSTYRQKMOLEQMLTALDOMRSIVSELAGLISAMEYOKTLTDEEL 240
QY 241 ADMKRREPIACIGPPNICLDRLNEMITSLAESQLOTRQOIKLELQOKYSYKGDPIVQ 300
DB 241 ADMKRREPIACIGPPNICLDRLNEMITSLAESQLOTRQOIKLELQOKYSYKGDPIVQ 300
QY 301 HRPMLERIVELFRNLMSAFVVERQPCMPHPRPLVKTGVQFTKVRLLVPRPELNY 360
DB 301 HRPMLERIVELFRNLMSAFVVERQPCMPHPRPLVKTGVQFTKVRLLVPRPELNY 360
QY 361 QLKIKVCDIKDSGVAAALRGSRKFNILGTNTKVMNMESSNNGSLSAEKKHILTRORCGN 420
DB 361 QLKIKVCDIKDSGVAAALRGSRKFNILGTNTKVMNMESSNNGSLSAEKKHILTRORCGN 420
QY 421 GGRANCASLIVTEELHITFETEVYHOGKIDLETSLPVAVSNICOMPANAFSTLAMY 480
DB 421 GGRANCASLIVTEELHITFETEVYHOGKIDLETSLPVAVSNICOMPANAFSTLAMY 480
QY 481 NMLTNKRKNVFFTKPPGTMDOVAEVLVSWQFSSSTTKRGLSIEOLTTLAEKLLGQVAYS 540
DB 481 NMLTNKRKNVFFTKPPGTMDOVAEVLVSWQFSSSTTKRGLSIEOLTTLAEKLLGQVAYS 540
QY 541 GCOITMAKFCENNAKGKFSFWMLDNIIDLVKYIYALAMGEGYMGTSKEREALIST 600
DB 541 GCOITMAKFCENNAKGKFSFWMLDNIIDLVKYIYALAMGEGYMGTSKEREALIST 600
QY 601 KPPGTFLLRSESSKEGVTFTWEKDISGKTQIOSVEPYTKOOLNNNSFEEIIMGYKIM 660
DB 601 KPPGTFLLRSESSKEGVTFTWEKDISGKTQIOSVEPYTKOOLNNNSFEEIIMGYKIM 660

DB 601 KPPGTFLLRSESSKEGVTFTWEKDISGKTQIOSVEPYTKOOLNNNSFEEIIMGYKIM 660
QY 661 DATNLIVSPFLVLYYDPIPEEAFGKYCRPESQEHPEADGSAAPYLKTFICVFTPTCSN 720
DB 661 DATNLIVSPFLVLYYDPIPEEAFGKYCRPESQEHPEADGSAAPYLKTFICVFTPTCSN 720
QY 721 TIDLPMSPRTLDLMQFGNNGEAEPSACGPFESLTFDMDLTSECATSPM 770
DB 721 TIDLPMSPRTLDLMQFGNNGEAEPSACGPFESLTFDMDLTSECATSPM 770
RESULT 4
STAI_HUMAN STANDARD: PRT: 750 AA.
AC P42224;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1-ALPHA/BETA
DE (TRANSCRIPTION FACTOR ISGF-3 COMPONENTS P91/P84).
GN STAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 514-524; 654-660 AND 667-672.
RX MEDLINE=92366557; PubMed=1502203;
RA Schindler C., Fu X.-Y., Impirota T., Aebersold R., Darnell J.E. Jr.;
RT "Protein of transcription factor ISGF-3; one gene encodes the 91 and
RL 84-kDa ISGF-3 proteins that are activated by interferon alpha";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7836-7839(1992).
RN [2]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RA Yan R., Qureshi S., Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in
RL coincident sites in STAT1 and STAT2";
RL Nucleic Acids Res. 23:459-463(1995).
RN [3]
RP PHOSPHORYLATION OF TYR-701.
RX MEDLINE=95386533; PubMed=7657660;
RA Quella F.W., Thierfelder W., Wiltuhn B.A., Tang B., Cohen S.,
RA Thie J.N.;
RT "Phosphorylation and activation of the DNA binding activity of
RT purified STAT by the Janus protein-tyrosine kinases and the
RL epidermal growth factor receptor";
RL J. Biol. Chem. 270:20775-20780(1995).
RN [4]
RP PHOSPHORYLATION OF SER-727, AND MUTAGENESIS.
RX MEDLINE=95354205; PubMed=7543024;
RA Wen Z., Zhong Z., Darnell J.E. Jr.;
RT "Maximal activation of transcription by STAT1 and STAT3 requires both
RT tyrosine and serine phosphorylation";
RL Cell 82:241-250(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 136-710.
RX MEDLINE=98292180; PubMed=9630226;
RA Chen X., Vinkemeier U., Zhao Y., Jeruzalmi D., Darnell J.E. Jr.,
RA Kurlyan J.;
RT "Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to
RT DNA";
RL Cell 93:827-839(1998).
CC -I- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -I- PATHWAY: IFN-SIGNALING PATHWAY
CC -I- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT. IN
CC RESPONSE TO IFN GAMMA, STAT1 FORMS HOMODIMERS, THAT ALSO

CC		TRANSLocate INTO THE NUCLEUS TO ACTIVATE IFN GAMMA-RESPONSIVE GENES. INTERACTS WITH NML.
CC	-I-	SUBCELLULAR LOCATION: NUCLEAR; TRANSLocATED INTO THE NUCLEUS IN RESPONSE TO PHOSPHORYLATION.
CC	-I-	ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA/p91 (SHOWN HERE) AND BETA/p84; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC	-I-	PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-GAMMA, IFN-ALPHA, PDGF, AND EGF. SERINE PHOSPHORYLATION IS ALSO REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (LACKING IN BETA FORM).
CC	-I-	SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC	-I-	SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outposton -- the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC		-----
DR	EMBL:	M97935; AAB64012.1; -;
DR	EMBL:	M97936; -; NOT_ANNOTATED_CDS.
DR	EMBL:	U18662; -; NOT_ANNOTATED_CDS.
DR	EMBL:	U18663; -; NOT_ANNOTATED_CDS.
DR	EMBL:	U18664; -; NOT_ANNOTATED_CDS.
DR	EMBL:	U18665; -; NOT_ANNOTATED_CDS.
DR	EMBL:	U18666; -; NOT_ANNOTATED_CDS.
DR	EMBL:	U18667; -; NOT_ANNOTATED_CDS.
DR	EMBL:	U18668; -; NOT_ANNOTATED_CDS.
DR	EMBL:	U18669; -; NOT_ANNOTATED_CDS.
DR	EMBL:	U18670; -; NOT_ANNOTATED_CDS.
DR	PDB:	1BF5; 12-AUG-98.
DR	TRANSFAC:	T01492; -;
DR	MIM:	600555; -;
DR	InterPro:	IPIR000980; SH2.
DR	InterPro:	IPIR001217; STAT.
DR	Pfam:	PF00017; SH2; 1.
DR	Pfam:	PF001017; SPAT; 1.
DR	SMART:	SMO0252; SH2; 1.
DR	PROSITE:	PS50001; SH2; 1.
KW		Transcription regulation; DNA-binding; Nuclear protein;
KW		Phosphorylation; SH2 domain; Alternative splicing; 3D-structure.
FT	DOMAIN	573 .. 670
FT	MOD_RES	701 .. 701
FT	MOD_RES	727 .. 727
FT	VARSPLIC	713 .. 750
FT	MUTAGEN	727 .. 727
FT		
FT	CONFLICT	261 .. 271
QO	SEQUENCE	750 AA; 87334 MW; 054A813522364BA6 CRC64;

Query Match	50.1%	Score 2017:	DB 1:	Length 750:
Best Local Similarity	52.5%	Pred. No.	5.8e-119:	
Matches 389;	Conservative 145;	Mismatches 193;	Indels 14;	Gaps 9;
QY	1	MAONNLOOLDPTRYAKOHLQHYSDFFPEMLKQFLPWIESODMAYAAASKESATLVFNL	60	
:	:	: :	:	:
Dd	1	MSOWEYOOLDSKFLEOHYHLYDPSFMEIRIOTYLAWLMEKDQWDEHAANDVSFATIRFDL	60	
QY	61	LGEIDDOYSRFLQESNNVLYQHNRIRIKQFLQSRYIEKPMETARIYARCLMESRLQTAA	120	
:	:	: :	:	:
Dd	61	LSQDDDOYSRFSLENNPLFHÖHNIRKSKRNLDQNFEDDIOMSMIIYSCLKBEERKLLENAQ	120	
QY	121	TAAQGGGAQNHPTAAVTEETEKOQMLEOHLDADRYKRPODEIQKQKVENLQDFDEFNYTK	180	
:	:	: :	:	:
Dd	121	RFNQ--AASGNIOSTVMLDKQEKELDSKVBNVKDWCMICEHEIKSELDQDEYDKCKLT-	177	
QY	181	SQGMDQLNGNNOASTTRKQMQLDEMLTALDQMRISYISELAGLLSAMEYQVKTLTDEEL	240	
:	:	: :	:	:
Dd	178	-QNREHNEINGAKSSQOKOEOLTKKKMYMTLNLRKREVHHKITELLNVTELTONALINDEL	236	
QY	241	ADMKRRPIACIGGPNICDLRLNMWITSLAESQJQTFROQIKKLEELDQKYSGDPVQ	300	

[illegible]

RESULT	5
STAL_MOUSE	STANDARD: PRT: 749 AA.
ID	STAL_MOUSE
AC	P42225:
DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	20-AUG-2001 (Rel. 40, Last annotation update)
DE	SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 1.
GN	STAT1.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eumetazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID: 10090;
RN	[1]
RP	SEQUENCE: A.
RC	TISSUE=Th;
RX	MEDLINE=1974; PubMed=7545930;
RA	Zhong Z., Wen Z., Darnell J.E. Jr.;
RT	"Stat3 and Stat1: members of the family of signal transducers and
RT	activators of transcription.";
RL	Proc. Natl. Acad. Sci. U.S.A. 91:4806-4810(1994).
CC	-1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC	RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC	TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC	-1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC	ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
CC	TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
CC	TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
CC	SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT. IN
CC	RESPONSE TO IFN GAMMA, STAT1 FORMS HOMODIMERS, THAT ALSO
CC	TRANSLOCATE INTO THE NUCLEUS TO ACTIVATE IFN GAMMA-RESPONSIVE
CC	GENES. INTERACTS WITH NFKB1 (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC	RESPONSE TO PHOSPHORYLATION.
CC	-1- INDUCTION: BY IFN AND EGF.
CC	-1- PPM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-GAMMA, IFN-ALPHA,
CC	PGF2, AND EGF. SERINE PHOSPHORYLATION IS ALSO REPORTED FOR MAXIMUM

```

CC      TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC      -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
DR      EMBL: P06924; AAA19454.1; -.
DR      HSP: P42224; 1BF5.
DR      TRANSFAC: T01575; -.
DR      MGD: MGI:103063; Stat1.
DR      InterPro: IPR000980; SH2.
DR      InterPro: IPR001217; STAT.
DR      Pfam: PF00017; SH2; 1.
DR      Pfam: PF01017; STAT; 1.
DR      SMART: SM00252; SH2; 1.
DR      PROSITE: PS00001; SH2; 1.
KW      Transcription regulation; DNA-binding; Nuclear protein; SH2 domain;
KW      Phosphorylation.
KW      DOMAIN 573 670 SH2.
FT      MOD_RES 701 701 PHOSPHORYLATION (BY JAKS) (BY
FT      MOD_RES 701 701 SIMILARITY).
FT      MOD_RES 701 701 PHOSPHORYLATION (BY SIMILARITY).
SQ      MOD_RES 727 727 PHOSPHORYLATION (BY SIMILARITY).
      SEQUENCE 749 AA: 87197 MW: 249D919952B65F1 CNG64;

Query Match 49.2%; Score 1984; DB 1; Length 749;
Best Local Similarity 51.0%; Pred. No. 6,8e-117;
Matches 390; Conservative 141; Mismatches 205; Indels 28; Gaps 11;

OY 1 MAOWMLOGLDTRRYKOLHOLYSDTFPMPLRQFLPWEESODMAYAAKRESIATLVFNL 60
DB 1 MSQWFLQDDLSFQYVHQLYDPSMPTROYLKLQWLEKQMEHAADVASATIRFDL 60
OY 61 LGEIDQVSRFLQESNVLAQHNLRKIKQFLQSRYLEKPMETARIYARGLWESESLRFL 120
DB 61 LSQLDDQVSRFLQESNVLAQHNLRKIKRKRNLQDNFQEDPVQSMIYNLQLEKERKLL 120
OY 121 --TAAGGQGANHPRAAVYTEKQMLDEGLDYKRRVODLEQKMKVVENLDDPFPNTKT 178
DB 121 RENQAOEGNIQ---TWMLDKQKELDSKVRNVKQDVNCIDELKLTLELDEYDFKCT 176
OY 179 LKSGQMDQILNNGNSVTRQKMOLEOMLTALDQRRSIVSELAELTAMEYVQKTLDE 238
DB 177 --SQNRGSLANGVANSQDQKQEBOLLHKHFMFLMDNKRKELIKIRLMSIELQWTLND 234
OY 239 ELADMKRRELICIGPPNICLDRLNNMTLSAESQLOTROUIKKLELQOKVSYGKDP 298
DB 235 ELVEVMKRQOASCIGPPNACIDQLOTQFTTVALELQIROOLKLELEQKFTYEPDI 294
OY 299 VQHRPMLERLYELFRNLKMSAFVVERQPCMHDPRLVKTGQFTKYRVLVFPPEL 358
DB 295 TKNQVLSRFTFLRQOLQSSFYVERQPCMPTHQRPLVKTGQFTKYRVLVFPPEL 354
OY 359 NYOLKIKVCIDKSDVAALRGSRFFNLGTNTKYMNEESNGSLSAEFLHLLREORC 418
DB 355 NLTFTVCKCFHFDVDEKMTVKGFRFFNLGTHTKYMNEESTNGSLAEFLHLLQLEKQK 413
OY 419 GNGGKANDCASLITFEEELHITFEFEYVHOGKILDLFHSI.PYVVISICQPMANASIL 478
DB 414 -NAGKRTNEGPLIVTELSHSFETQLQCPGLVIDLETTTSPVAVVISVNSQLPSGASIL 472
OY 479 WYMLJLNNKKNVNFETKPPIGTMDVAEVLWSQFSSTTKRGLSIEQLTTLAEKLGSPVN 538
DB 473 WYMLJLNNKKNVNFETKPPIGTMDVAEVLWSQFSSTTKRGLSIEQLTTLAEKLGSPVN 532
OY 539 YSGCQITAAKPKCKENNAKGFSTFWYVLDNIIDLYKKYTLALMNGCYIMGFTSKERRAIL 598
DB 539 YSGCQITAAKPKCKENNAKGFSTFWYVLDNIIDLYKKYTLALMNGCYIMGFTSKERRAIL 598

```

```

Db 553 PDGL-1PMTRECKENINNDKNSFWHMDIILLLEIKINDLLCLMNDDCIMOFISKERERALL 591
Qy 599 STKPPGTELLARESSSKBEGCVTFVWEKDIS-GKTOIQSVEPYTQOQLNNMSFAELIMGY 657
Db 592 KDOQGCTFLLRSESSSRCAIIFTFVWERSQNGCEPDAFVPEYTKKELSAVFPFDIIRNY 651
Qy 658 KIMDPTNLIVSLVLYLPDIPKEAFAFGK-YCPR-ESQEHPEADPCSAALYTKFKVCV- 714
Db 652 KYMAENIENPENLKLTYLPIDKDHAFGKYYSRKPEAPEFMELDDPKRTGYITELISVSE 711
Qy 715 -PRTCSMTID-LPMSPTLIDSLMFGNGNDEGAERSAGQFESL 755
Db 712 VAPSRLQTTDNLPMSPSEFDEMSRT-----VGPEDFSM 745

RESULT 6
STA4_HUMAN STA4_HUMAN STANDARD: PRT; 748 AA.
AC Q14765;
DT 30-MAY-2000 (Rel. 39, created)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 30-MAY-2000 (Rel. 39, last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 4.
GN STA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu X., Sun Y.L., Hoey T.;
RT "The STAT amino-terminal domain mediates cooperative DNA binding
RT and confers selective sequence recognition."
RL Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION.
CC -!- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -!- PTM: TYROSINE PHOSPHORYLATED. SERINE PHOSPHORYLATION IS ALSO
CC REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L78440; AAB05605.1; -.
DR MIM: 600558; -.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SMO0252; SH2; 1.
DR PROSITE: PSS0001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain.
FT DOMAIN 569 664
FT MOD_RES 693 693 SH2.
FT FT PHOSPHORYLATION (BY JAKS) (BY
FT FT SIMILARITY).
MOD_RES 721 721 PHOSPHORYLATION (BY JAKS) (BY
SEQUENCE 748 AA; 85940 MW; 11E43803A9A4FFA CRC64;

Query Match 44.6%; Score 1798; DB 1; Length 748;
Best Local Similarity 48.4%; Pred. No. 3e-105;
Matches 357; Conservative 146; Mismatches 209; Indels 26; Gaps 10;

```

```

OY 1 MAOMNOLQDLTRYKQLHOLYSDTFPMELOFLAPMIESQDMAVAASKESHATLVFNHL 60
   1 MSQMNVOQOLEIKFLBOVDQFDNFPMEIRHLLAQMIENQDWEASNNMTATILLQNL 60
OY 61 LGETIDQYKRFLOESNVLYOHNLRIKQLOSLRYLEKPMETIARIVARCLMEESRLQTA 120
   61 L1QLDLQGLGVSKENKLLHNLKIRIKVLYQGFHGNPMHVAIVAVVNSCLREBRIL-AAA 119
OY 121 TAAQGGQANHPAAVYTEKQMLFQHLQDVRKRYODLEQMKVENVLDDDFPNYKTLK 180
   120 NMPVQGLEKSLQSSVSEQDVEHKAIVAKNSVQMTEDDTYLEDLODFDYRYKTLTQ 179
OY 181 SOGDMODLNGNOSVTRQKMOLEQMLTALDQRRSIVSELAQLLSAMEYVQKTLDEEL 240
   180 T---MDQSDKNSAMVNOEVLTLQEMLNLSLDFEKRKALSMTQIIHETDLMTMLLEEL 235
OY 241 ADKRRPELACIGGPNICIDRLNNITSLAESQLOTRQOIKKLELEQOQVSKGDPYQ 300
   236 QDKRRQOILACIGGPLHNGLDLQNCFTLLAESLFLQRLKLELEQDSTMTTEGDIYPM 295
OY 301 HRPMLEERIVELFRNLKMSAFVVEROPCMHPRDPLVITGVQFTTKVRLKFPPELNY 360
   296 QRTHMLERYVFLYNLFKNSFVEVEROPCMPTHPQRPVLTKLQFTYKRLLIKIPELNY 355
OY 361 QLKIKVCIDKSDVAALKSKRFNLGNTKYNNNEESNGSLSAEFKLLTLREORCGN 420
   356 QVYKVASIDK---NVSTL-SNRPFVLCGNVVKAMSIEESSNGSLSEFRRHLOREKMSA 411
OY 421 GGRANDASLIYVEELHLLFEFEVYHOGKIDLETHSLVAVVVISNCOMPMASITLWY 480
   412 GGGGN-EGCMAMTVEELHSTTFETQICLYGLTIDLETSSLPVVISNVSQLPNMASTIY 470
OY 481 NMLTNPKNVNFFTPKPIGTMDQVAEVLWSQFSSTTKRGLSIBQLTTLAKLLGPVNS 540
   471 NVSTNDSQNLVFFNNPBPALISQLEVMWSQFSYVGRGLNSDQLHMLAKLTQVOS-SYS 529
OY 541 GCQITAKCKEMAKGPFSPFWMLDNIIDLVKYLALNEGTYGFIKKEBRALST 600
   530 DGHLTAKKFKHEHLPKSPFTFMTWLEALIDLKKNHLLPLMIDYVGVGFSKEXERLLKD 589
OY 601 KPPGTFLRFSESSKGGVTFWWEKDISGKTQIQSVEPYTKOOLNMSFAETIIMGYKIM 660
   590 KMGFTFLRSESHL-GGIFTWVDHSEGEVHFHSEVPKRGRLSALPRADILLROKVI 648
OY 661 DATNIIYSLPYLYPDIKPEAKGKYCRPE---SOEHPADGSAAPYKTKFICVPTPT 716
   649 MAENLPEENLPKLYLPDIKPKDAFGKHYSSQPCVSRPTERGDKG---YVPVFIPISTI 704
OY 717 TCSNTLI-----LPMSP 728
   705 RSDSTEPHSPDLPMSP 722

```

```

RT activators of transcription.*;
RN Proc. Natl. Acad. Sci. U.S.A. 91:4806-4810(1994).
RN [12]
RN SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-Testis;
RX MEDLINE-94277038; PubMed-8007943;
RA Yamamoto K., Quelle F.W., Thierfelder W.E., Kreider B.L.,
RA Gilbert D.J., Jenkins N.A., Copeland N.G., Silvennoinen O.,
RA Ihle J.N.;
RT Stat4, a novel gamma interferon activation site-binding protein
RT expressed in early myeloid differentiation.*;
RT Mol. Cell. Biol. 14:4342-4349(1994).
RN [13]
RN X-RAY CRYSTALLOGRAPHY (1.45 ANGSTROMS) OF 1-123.
RX MEDLINE-98128033; PubMed-9461439;
RA Vinkemeier U., Moareff I., Darnell J.E., Jr., Kurian J.;
RT Structure of the amino-terminal protein interaction domain of
RT STAT-4.*;
RL Science 279:1048-1052(1998).
CC -1- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
CC ACTIVATION OF TRANSCRIPTION.
CC -1- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
CC MEMBER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- PTM: THROSONE PHOSPHORYLATED. SERINE PHOSPHORYLATION IS ALSO
CC REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 006923; AAA19453.1; -.
DR EMBL: 009351; AAA19692.1; -.
DR PDB: 1BGF; -SEP-98.
DR TRANSFAC: J01576; -.
DR MGD: MG1J04062; Stat4.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00...42; 1.
DR Pfam: PF01...FAT; 1.
DR SMART: SM...SH2; 1.
DR PROSITE: P50001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain; 3D-structure.
FT DOMAIN 570 665
FT MOD_RES 694 694
FT PHOSPHORYLATION (4*) (K)S (BY
FT SIMILARITY).
FT MOD_RES 722 722
FT PHOSPHORYLATION (4*) (S) (BY
FT CONFLICT 198 638 MISSING (IN REF. 2).
FT CONFLICT 198 638 A -> P (IN REF. 2).
SQ SEQUENCE 143 AA; 85940 MW; A88B837E49CFEBC CRC64;

Query Match 44.2%; Score 1780.5; DB 1; Length 749;
Best Local Similarity 48.1%; Pred. No. 3.8e-104;
Matches 356; Conservative 149; Mismatches 206; Indels 29; Gaps 11;

```

```

Db 120 NMPDQPLEKSSVSEORVNEKVAIKSVQMTEDDTYLYEDLOEDFYRYKTIQ 179
Oy 181 S--QGMODLNGNNSVTRQKMOOLEMTALDPMRRIYSELGLLSAMEYQKTLTDE 238
Db 180 TMDQGRK-----NSILVNEQVLTLLDQMLNSLDPKRKREALKSKMTQIVNEFDLMSMLE 234
Oy 239 ELADWKRREPLACIGPPNLCIDRLNEMWITSAESLOLQTRQIKLEELQOKVSKDPI 298
Db 235 ELQDMKRRQOICIGPLNGLDQIONCFLLAESL.FOLQOLEKLEQOSTKMTMYEGDPI 294
Oy 299 VOHRPMLERIEVLEFNLMKSAFVNERQCPMPHPPRPLVYIKGVQFTTKVRLVAFPEL 358
Db 295 PAORALLERAFLLYNLFRKNSFVVERQPCMPHPORPMVYKTLQFTVYRLRLIKLPEL 354
Oy 359 NPOLKIKVCIKDSQDVAAALRGSRKFNILGTNTKVNMMESNNGSLSAEFKHLTLREORC 418
Db 355 NVOYKAKASIDK---NVSTL-SNRREVLCGTIHKAMSSSESSNGSLSYVERHLQPKEMKC 410
Oy 419 GNGGRANCASLIVTEELHLITFEVEYHOGKIDLETHSLPVVVISNICQMPNANASIL 478
Db 411 STGSKGN-EGCHVTEELSHSTFETOICLYGLTINLETSSLPVVVISNVSQLPNANASII 469
Oy 479 KVNMLLNKKNVNEFTKPRIGTMDQVAVELSMQSSSTTKRGLSEDTTLAEKLLRGVN 538
Db 470 WYNVSTNDSONLFEFNPPSVYTLGQLEVMASWQFSSYVGGLNSEQDLMNAELKLVQVS-N 528
Oy 539 YSCQITMAKFKCKENNAAGCFSPWVLDNIDIVKRYILALMEGYIMGPIISKEREAIL 598
Db 529 YNDGHLTMAKFKCEHLPGKTFETFWLLEAILDLIKKHILPLMDGYIMGVSKEREKLL 588
Oy 599 STKRPCTFLLRSESKSEKGVFTTWYKEDIQSQTOIQSEVPYTKOOLNNMSPAEIIMGYK 658
Db 589 KKMPEFTFLLRSESHL-GGIFTWVDOSENGVRFHSEVPYKGRISALAFADIIIRDYK 647
Oy 659 IMDATNITLSPYLYLPDIPKPEAFGKYCRPE-----SOEHPDPSAAYLTATKFCVY 714
Db 648 VIMAEINPENPLKYLTPDIPKKAFAKHHISSQCEVSRPERDCKG-----YVPSVFILPIS 703
Oy 715 PPTGSMFTID-----LPMSF 728
Db 704 TIRSDTEPQSPDLPMSP 723

RESULT 8
STA2_HUMAN STANDARD: PRT: 851 AA.
AC P52630: Q16430: Q16431:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2 (P113).
GN STAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95192056; PubMed=7885841;
RA Yan R., Qureshi S., Zhong Z., Wen Z., Darnell J.E. Jr.;
RT "The genomic structure of the STAT genes: multiple exons in
RT coincident sites in Stat1 and Stat2.";
RL Nucleic Acids Res. 23:459-463(1995).
RN [2]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=93366558; PubMed=1502204;
RA Fu X.-Y., Schindler C., Impirota T., Abersold R., Darnell J.E. Jr.;
RT "The proteins of ISGF-3, the interferon alpha-induced transcriptional
RT activator, define a gene family involved in signal transduction.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7840-7843(1992).
RN [3]
RP ALTERNATIVE SPLICING.

```

```

RX MEDLINE=96176320; PubMed=8601453;
RA Sugiyama T., Nishio Y., Kishimoto T., Akita S.;
RT "Identification of alternative splicing form of Stat2.";
RL FEBS Lett. 381:191-194(1996).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=97172544; PubMed=9020188;
RA Bluyssen H.A., Levy D.E.;
RT "Stat2 is a transcriptional activator that requires sequence-specific
RT contacts provided by stat1 and p48 for stable interaction with DNA.";
RL J. Biol. Chem. 272:4600-4605(1997)
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA.
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U18671; AAA98760.1; -
DR EMBL: M97934; -; NOT_ANNOTATED_CDS.
DR EMBL: S81491; AAB36226.1; -
DR EMBL: S81491; AAB36227.1; ALT_SEQ.
DR HSSP: P42224; 1BF5.
DR MIM: 600556; -
DR InterPro: IPR000980; SH2
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain; Alternative splicing.
FT DOMAIN 572 667
FT MOD_RES 690 690
FT VARSPPLIC 621 652
FT FT
FT FT
FT FT
FT VARSPPLIC 653 851
FT FT
FT FT
FT VARSPPLIC 851 851
FT FT
FT FT
FT SEQUENCE 851 AA; 97916 MW; E4C74674CB7A3215 CRC64;

```

```

Query Match 32.1%; Score 1293; DB 1: Length 851;
Best Local Similarity 38.6%; Pred. No. 166-73;
Matches 298; Conservative 148; Mismatches 276; Indels 50; Gaps 18;

Oy 1 MAQWNLQQLDTRYLKQLHQLVSDT-FPMELRQFLAPWIESQDMAYAA--SKESHATLVF 57
Db 1 MAQWEMQLNDSPFOQLHQLVSHSLPVDIRQLAVWIEDQWQGAALGSDSKATIMF 60
Oy 58 HNLGELIDQYSFLOE-SVLYQVHMLRKQFLQSRYLEKPMELARIYARCLWESRLL 116
Db 61 FHLDLQNLVCEGCSQDPSLQLDHLNLRKFCRDIDP-FSQDPTQLAEMIFNLLEKRIL 119
Oy 117 QTAATV AVOGGGANHPETAAVVTEKQO-MLDQHLQDVRRKRVQDLEQMKVVENLDQDFDN 175

```



```
Db 120 IQAORAEQGE---PULETFVESQOHEIESRIIDLRAAMEKLYKSISQLODQDQVFCFR 176
Qy 176 YKTLKSGQDMODLNGNNSVTRKMOOLEOMLTALDQMRISIVSELAGLSAMEYVQATL 235
Db 177 YK-IQANGKSTSLDPH--QTEQEKI--LQETLNELDKRRKREVDLASKLGLGTLTLEIEL 231
Qy 236 TDELDAMKRRPELIACIGGPPNITCLDLENNITSLAESQLOTRQOIKLELEQKLVSKG 295
Db 232 L-PLEEMKRAQOQKACIRAPIDHLEOLETFETAGAKLLFHLROLKLEKLSCLSVSYOD 290
Qy 296 DPLVQHHPMLEERIVLEFRNLKMSAFYVEROPCMRPHDRPLVITKGTQFTTKVALLKF 355
Db 291 DPLKGVDLRAHQYTLEQLRLHRAFYVETOPCMQTPHRLILKTKGSKTYRRLVRL 350
Qy 356 PELVYQLIKVCIDKDSQVVALRGSRKFNILGTNTKVMNESNGSLSAEFKHLTRE 415
Db 351 QEGNESLVEVSIIDRNP---POLQGFRRFNILTSNOKLTLPREKQSGQLINDCYLTVE 407
Qy 416 QRCNGGGRANDASLIYTEELHITFETEVYHOGKIDLETHSLPYVYVINSICMPNMA 475
Db 408 QRSQSGSGSKKCPGLGYTEELHISFYVKYTYQGLKQELKDTLPVYIISNNQSLIAMA 467
Qy 476 SILVYNNMLTNNKVNPFETKPPITGTDQVAEVLWSQFSSTTKRGLSIDQTLTLEKLG 535
Db 468 SVLFNPLSLPRLQNOQFFSNPKAPMSLLGPALSMQFSSYVGRGLNSQSLNRKRLKG 527
Qy 536 GVNYSQCOITWAKFCR-ENMAGKGFSEVWMLDNIIDLYKKITLALWEGYIMGPTISKRE 594
Db 528 NCRTEDPLTSLVADPTKRSPPGK-LPEFTWLDKILLELVHDLKLELNMGRINGEFSRSQ 586
Qy 595 RAISTKPPGTFLEFRFSSSEKGGVTFWVEKIDSKTQIOSVEPYTQOOLNMSFALI 654
Db 587 RRLKTKMSGTFLRFSSS-EGGITCSWVHODDKVLYISVOYTKVEVLSQSLPTLEII 645
Qy 655 MGXKIMDATNLLVSPLYLPDIKEAFAKRYCRPE--SQEHPEADPGSAAPYLTKTFC 712
Db 646 RHYQLLEENIPENRPLRPLRIPRIDEAFGCGYQEKVNLQERRK-----YLKRLIIV 697
Qy 713 VTPITCSNTIDLPMSPRLDSLMQPGNNGEAFSAGQFSLTFTMDLTSE 764
Db 698 V-----SNQVDELOQ-----PLELKPEPELESLELELGLVPE 730

RESULT 9
STA2_MOUSE STANDARD; PRT; 923 AA.
AC G9WVL2; Q64189; Q64250; Q64188;
UT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RX 111
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1;
RA Paulson M.S., Mui A., Levy D.E.;
RT "Molecular cloning and characterization of murine Stat2.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN 121
RS SEQUENCE OF 595-658 FROM N.A. (ISOFORMS A AND B/C).
RA MEDLINE=96176320; PubMed=8601453;
RA Sugiyama T., Nishio Y., Kishimoto T., Akira S.;
RT "Identification of alternative splicing form of Stat2.";
RL FEBS Lett. 381:191-194 (1996).
CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
CC TRANSCRIPTION FACTOR IS TERMED ISGF3.
CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
```

```
CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B/C. ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BRAIN, LUNG, HEART, SPLEEN,
CC LIVER, KIDNEY, MUSCLE, AND THE TESTIS.
CC -1- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF088862; AAD38329.1; -
DR EMBL: S81342; AAB36228.2; -
DR EMBL: S81342; AAB36231.1; -
DR EMBL: S81342; AAB36230.1; ALT_SEQ.
DR HSSP: P42224; 1BF5.
DR MGD: MGI:103039; Stat2.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001217; STAT.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF01017; STAT; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Phosphorylation; SH2 domain; Alternative splicing.
FT DOMAIN 571 666
FT MOD_RES 689
FT VARSPPLIC 689 689 PHOSPHORYLATION (BY JAKS) (BY
FT SIMILARITY).
FT FT HKEIYSQPTKREVIQSLPTREI -> GQHPYHSGSL.
FT SARHPDRLPL (IN SHORT ISOFORM).
FT VARSPPLIC 923 923
FT FT MISSING (IN SHORT ISOFORM).
FT CONFLICT 596 596 T -> A (IN REF. 2).
FT FT 620 620 H -> D (IN REF. 2).
SQ SEQUENCE 923 AA; 105416 MW; D50BB54C535H0774 CRC64;

Query Match 31.6%; Score 1271.5; DB 1; Length 923;
Best Local Similarity 38.8%; Pred. No. 4e-72;
Matches 292; Conservative 133; Mismatches 265; Indels 63; Gaps 18;

Qy 1 MAONNQLOLDTRYKJOLHOLYSDFP-PMEILROPLAPMIESQDMAYAA--SKESHATLVF 57
Db 1 MAQWEMKLNLSPPDLQLOVYSQSFLPMDFRQHLASWIEDQNMREDALESQDAANMLY 60
Qy 58 HNLGEIDQOYSRFLQESNVLYOHNLIRIKQFLQSGRYLEKPMELIARIARCLMEERSRLQ 117
Db 61 FSLIDQLQNMWHYSSDPRSLILQHLNRKFSRDIOF-FPNGSPQALAMIFNLLEQRIL- 118
Qy 118 TAATTAQGGGGAANHPTA--AVYTEKQMLEOHLDVRRKRVODLEQKMKVENLQDDPDFN 175
Db 119 ---IQAORAEQVOPAPAEAVESQOLEIENRICQIHLVHDIETFLVRSIRQLKEDQVFSFR 175
Qy 176 YKTLKSGQDMODLNGNNSVTRKMOOLEOMLTALDQMRISIVSELAGLSAMEYVQATL 235
Db 176 YTVF-----SLKTTSSDPHOSQOAHVVOATAKVPMRREVDLISKGLVGRV---T 224
Qy 224 TLTD---ELADWKRKRPETACIGGPPNI--CLDRLENNITSLAESQLOTRQOIKLELEQ 288
Db 225 TLVDLLPKLDEBWKVQOQKSCIGAPPVYKSAAEQLEQMLTGAKAKLFLRLQRLQKLEM- 283
Qy 289 QKVSIGDPIVQHRPMLEERIVLEFRNLKMSAFYVEROPCMRPHDRPLVITKGTQFTTK 348
```

```

Cc Db 284 SCLRTQGDMEFAKGVLDLRNAQVWELLRLQRLQSRFSVETQPCMPQTLRLPLTLTKGNKFTVR 343
Cc Qy 349 VLLVYKFPPELNYOLKIKVCIIDKSDQVAALRGSRKFNILGTITKVMNMESSNNGSLAEF 408
Cc Db 344 TRLVLRLEQSGESLSLAEVSDRNSD---LPQFRKFNILTSNQKTLTPKGGROGLIMDF 399
Cc Qy 409 KHLTLREQRCNGGRANCASLIYVEELHLITFEFVYHQGLKIDLETHSLPVVYISNIC 468
Cc Db 400 GFLTLVEORAVGAGKGNNGKPLAVEELHIVISFVEYTGQKMKLQTDTLTPVYIISNN 459
Cc Qy 469 QPNPMASITLWYNNMLTNPNKRVNFTKPKPIGWQDAEVLWSQFSSTTKRGSLIEQLTL 528
Cc Db 460 QLSFMASTLWNNMLSPNKNQOFPQCARKAWMSLLGPTLVSSFFSYVARGLDSGLML 519
Cc Qy 529 AEKLLGPGVYNSGCOITWAKPCK-ENNAKGFSFWYVLDNIIDLYKKYLLALNCEYIMG 587
Cc Db 520 RRLKLGKSCMKMEDALLSWPDFCKRESPPCK-TPFWTWLDKILELVHDHLKMDKGRIMG 578
Cc Qy 588 FTSKERERAILSTKPGFTLLRFSSSKRGQYTFWVEKDISGKQIOSEVETPQKQOLNN 647
Cc Db 579 FVSRNQRERLLKMLSGTLTLRFSEIS-BGGLTCSWEVHODDHKVEITYSVQPTKLVLOS 637
Cc Qy 648 MSFAELIMCYKIMDATNITLVSPLYLYDPIKEAEAFCKYCRP---ESQEHPEADPGSAA 703
Cc Db 638 LPLTEIIRHYOYLAENINPENLRFPLYPRIPEDEAFGCYOEKVNLEQEE----- 688
Cc Qy 704 PLYLTKFCVPTPTCSNTIDLPMSPRTDLSMW 736
Cc Db 689 -YLKHLKIYI-----SNRQVDELQ 707

Cc RESULT 10
Cc STA2_PIG ID STA2_PIG STANDARD: PRT: 864 AA.
Cc AC 002739;
Cc DT 15-JUL-1999 (Rel. 38, Created)
Cc DT 15-JUL-1999 (Rel. 38, Last sequence update)
Cc DT 15-JUL-1999 (Rel. 38, Last annotation update)
Cc DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.
Cc OS STAT2.
Cc OS Sus scrofa (Pig)
Cc OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Cc OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
Cc OX NCBI_TaxID=9823;
Cc HP [1]
Cc RN SEQUENCE FROM N.A.
Cc RC TISSUE-Muscle;
Cc RA Ito Y., Mukawa S., Kobayashi E., Wada Y., Minezawa M.;
Cc RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
Cc CC -I- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED
Cc CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN
Cc CC TRANSCRIPTION FACTOR IS TERMED ISGF3 (BY SIMILARITY).
Cc CC -I- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-
Cc CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON
Cc CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX
Cc CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT
Cc CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT (BY
Cc CC SIMILARITY).
Cc CC -I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
Cc CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).
Cc CC -I- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA (BY
Cc CC SIMILARITY)
Cc CC -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
Cc CC -I- SIMILARITY: CONSTAINS 1 SH2 DOMAIN.
Cc CC -----
Cc CC This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc CC the European Bioinformatics Institute. There are no restrictions on its
Cc CC use by non-profit institutions as long as its content is in no way
Cc CC modified and this statement is not removed. Usage by and for commercial
Cc CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc CC or send an email to license@isb-sib.ch).

```

CC	EMBL: AB004061; BAA20332.1; -	Query Match	31.4%; Score 1264.5; DB 1; Length 864;
DR	InterPro: IPR000980; SH2.	Best Local Similarity	37.6%; Pred. No. 9,9e-72;
DR	InterPro: IPR001217; STRAT.	Matches	292; Conservative 149; Mismatches 279; Indels 57; Gaps 18;
DR	Pfam: PF00017; SH2; 1.		
DR	Pfam: PF01017; STRAT; 1.		
DR	SMART: SM00252; SH2; 1.		
DR	PROSITE: PS00001; SH2; 1.		
KW	Transcription regulation; DNA-binding; Nuclear protein;		
KW	Phosphorylation; SH2 domain.		
FT	DOMAIN 572 667		
FT	MOD_RES 572 690		
FT	SH2.		
FT	PHOSPHORYLATION (BY JAKS) (BY SIMILARITY)		
SO	SEQUENCE 864 AA: 99060 MW: 28033569 FS B520 CRC64:		
QY	1 MAQNHQDLODLRYLKQLHQLYSDN-FPEMLKQFLAPWIESQDWAYA--SKESHATVLF 57		
DB	1 MAQWMLQDLSDPFQDHLQLYSESLPLPDVQOYLAWLIEDQWQEAALGNDCKRMMLF 60		
QY	58 HNLLEIDQYSRFQESN-VLYQHNLRKQFLQSRLEYKEMETARIYAKCLMESSRL 116		
DB	61 FHFIDQLNDCRCGCODEPCLLDHNLKRFYDIAQ-IPQGTPLAEMLFYNLLEEKRL 119		
QY	117 QTAATAAQQGGAHPYAAVTEKQ-MLEQHLQDYRKRYVDI EQRKRYVENLQDDFDEN 175		
DB	120 IQAQQAQL---EQQEPALAEAPGQDQHEIETSLLELRAMEKLIVKSIQOLKQDDIFCR 176		
QY	176 YKTLKSGQMDQLNNQNSVTRQKKQOLQQLTALDQMKRSIVSELAGLISAMEYVQKTL 235		
DB	177 YK-IKASAKYHSLDHR--TRQE-QVLOETNELDKRKEVLYDSKALIGKLTLLIEL 231		
QY	236 TDEELDAKKRREIACIGGPRNICIDRLNNMTTSLAESQLQTRQOIKKLEELQQKVSYKG 295		
DB	232 L-PRLEEMKVVQQAKCIGAPMDGELEQLEKMTAEKLLFHLKQLLEIKGSISSVKKYIE 290		
QY	296 DPVQHRPRLERIVELFRNLKSAFVYERQCPMHRDRLVITQVQFTTKVRLLYKV 355		
DB	291 DLEFGVDLLKAQVTELLQRLRLHRAFIYETQCPQTPRRPLRLTKGSKFYVTRLLVRL 350		
QY	356 PELNQLKIKVCIDKDSGDVAALRGSRFNLICNTNKVNMESNNNGSLISAEFKHLTYRE 415		
DB	351 QEGNESITAFVSTIDNP---PKSQGFRKFNILTSNRKITTPKGGSQGLIMPFGVLTLE 407		
QY	416 QRCGNGRANCDASLIVTEELHLITFETEVYHQGLKIDLETSHLPVYVNSITQCPNNA 475		
DB	408 QRSQAGGKNNKGPGLGVIETELHISFTVAKYVQGLKQELTDTPLPVIISNNQLS IAWA 467		
QY	476 SLIWMNMLTNNKNNNFTFKPRPIGTWQDVAEVLISWQSFSTRKGLSIDLTTLARKLGP 535		
DB	468 SLIWMNMLTNNKNNNFTFKPRPIGTWQDVAEVLISWQSFSTRKGLSIDLTTLARKLGP 527		
QY	536 GVNYSQCOITWAKFCF-ENMAGKGFSEFWWLDNI IDLVKKYLLALMNGEYIMGTFSKERE 594		
DB	528 NSSTGSLISWVDFIKRESPCK-LRFYMTWLDKIIDLVYHDHLKDLKMGKHIMGFVSRSE 586		
QY	595 RALSTKRPQGFLLAFSESSKEGVTFTWVEKDIGSKTQIQSVEPYTQOQLNNMSFAETI 654		
DB	567 RRLKKTITSGFFLRFSE-TLEGQITGCVWEHQDDDKVLYISLQPFTEKVEVLOSLPLTKII 645		
QY	655 MGYKIMDATNITSLVLYPDIPKEEAAGFKVCRESQDQHPRADGSAARYLKTFTICVT 714		
DB	646 SOYQLTLEENIPENLRLRLPRIPDEAFQCY---NDE--KANIQERKKYIKHLLIVAS 699		
QY	715 -----PTTCSNTIDLPMSPTIJDLSLMOFGNNGEGAEPSAGQFESL 755		
DB	700 NROYDELQAPPELKLLEPDLSELDLGLAP-----GPREGVDLDEPL 742		

	Query Match	21.5%	Score 667;	DB 1:	Length 799;
	Best Local Similarity	31.6%;	Pred. No. 7,1e-47;		
	Matches 240; Conservative	138;	Mismatches 303;	Indels 78;	Gaps 24;
<hr/>					
Oy	1 MAQNNQLDOLTRYIKLKHLYSDFFPHEMLROFLAPWIESODM----	AYAASKESHATLV	56		
	: :		:	:	
Dd	1 MAGMTIADLOOGDALRANOVKYLGHGFPLEVRVRYIAKWMSQPMADIDVDNPDQRAQAQL		60		
	:		:	:	
Oy	57 FHNLLGEIDQQYSRFLEQSNVLYQHNLRRIRKIQFLOSRLYLFKPMEIARIIVARCLWEESRL		116		
	: : : : : : :		:	:	
Dd	61 LFNVLQELKKRKAEHGVGEGGLFKLKTLCHAYATOLONTRDCRPMEIIVLRICIRILINVEOQL		120		
	: : : : : :		:	:	

RC STRAIN=C57BL/6A; A/J; 11550E-LIVER;
RX MEDLINE=95237198; PubMed=7720707;
RA Mul A. L. -F., Wakao H., O'Farrell A. -M., Harada N., Miyajima A.;
RT "interleukin-3, granulocyte-macrophage colony stimulating factor and
RT interleukin-5 transduce signals through two STATs homologs.";
RL EMBO J. 14:1166-1175(1995).
RN [2]
RP
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=96004632; PubMed=7568026;
RA Liu X., Robinson G. W., Gouilleux F., Groner B., Hennighausen L.;
RT "Cloning and expression of Stat5 and an additional homologue (Stat5b)
RT involved in prolactin signal transduction in mouse mammary tissue.";

RL Proc. Natl. Acad. Sci. U.S.A. 92:8831-8835(1995).

CC -i- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND

CC ACTIVATION OF TRANSCRIPTION. BINDS TO THE GAS ELEMENT AND

CC ACTIVATES PRL-INDUCED TRANSCRIPTION.

CC -i- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY

CC MEMBER (BY SIMILARITY).

CC -i- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN

CC RESPONSE TO PHOSPHORYLATION.

CC -i- TISSUE SPECIFICITY: IN THE VIRGIN, FOUND IN MOST TISSUES.

CC PARTICULARLY ABUNDANT IN MUSCLE TISSUE OF VIRGIN AND LACTATING

CC FEMALES, AND OF MALES.

CC -i- DEVELOPMENTAL STAGE: DETECTED BOTH IN VIRGIN MOUSE AND AFTER

CC MAMMARY GLAND INVOLUTION. THE LEVEL OF STAT5A INCREASES CONSTANTLY

CC DURING PREGNANCY, BUT DECREASES DURING LACTATION.

CC -i- P.TM: TYROSINE PHOSPHORYLATED.

CC -i- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.

CC -i- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL: 248539; CA88420.1; -

DR EMBL: 021110; AAC52282.1; -

DR HSSP: P42224; 1BP5.

DR TRANSFAC: T00944; -

DR MGD: MGI:103035; Stat5b.

DR InterPro: IPR000980; SH2.

DR InterPro: IPR001217; STAT.

DR Pfam: PF00017; SH2; 1.

DR Pfam: PF01017; STAT; 1.

DR SMART: SM00252; SH2; 1.

DR PROSITE: PS50001; SH2; 1.

DR Transcription regulation: DNA-binding: Nuclear protein:

KW phosphorylation; SH2 domain.

FT DOMAIN 589 686 SH2.

FT MOD_RES 699 699 PHOSPHORYLATION (BY JAKS) (BY

FT CONFLICT 433 433 SIMILARITY).

FT SEQUENCE 786 AA; 90002 MW; A8FE76405E41B2EF CRC64;

Query Match 21.5%; Score 865.5; DB 1; Length 786;

Best Local Similarity 30.1%; Pred. No. 8,6e-47;

Matches 247; Conservative 146; Mismatches 317; Indels 111; Gaps 30;

QY 1 MAOWNLOQLDTRRYLKQLHQLYSDPFPMELROFLAPWISQOMAYV-----ASKESHATLV 56

DB 1 MAMWIOAOOLQGDALHQMALYGOHEPIEVRYHLSQWISQAMDSIDLDPQENIKATOL 60

QY 57 FHNILGELIDQOYSRFLQESNVLYOHNLRIKQFLQSRYLEKPMELRIYARCLMESRL 116

DB 61 LGLVGLQKKAHEHVGEGFLKIKLGHYATOLQSTYRCMELVRCRLHILYNQORLY 120

QY 117 QTAATRAOOGGANHTAAVTEKQOMLEQHLQDVKKRVQDLQKKKVVENLQDDFFNY 176

DB 121 REANNSSSPAGS-----LADAMSQKHLOINQTEELRLITQDENELKKLQOQOEYFIQY 176

QY 177 -TILKSQGMQDLNGN-----OSVTRQKMOOLEQML-----TALDQRRSIVSELAL 224

DB 177 QESLRLQAOFOALQOLNPOERSRSTALQOKOVSLQETWLOREAQTLQOYRVELAEHQRT 236

QY 225 LSAMEYVQKTLTDEELADKKRRPELACIGPPNICIDRLNNITSLAESQLOTRQOIKL 284

DB 237 LQLLRKQGTIIILDELQKRRQQLANGNGPREGSLDVLOSMCEKLAELIIMNQROKIRA 296

QY 265 EELQGVKSTGKPIVOHRRPMEERIVELEFRNLKSAFVEROPCPMPHMDRPLVIRKTVQ 344

DB 297 EHLCOOLPIFG-PVEEMLAENVATITDLSALVTSTFIIEKQP-----FOVLKTOYRK 347

QY 345 FTTKRYELVFPPELNYL---KIKVCIDKDSGVAAALRSRKNF-----ILGTRKYVMN 396

DB 348 FAATVRLVG-KGLVHMPDPVKATITISPOAKSILKKNENTRNDYSGEIL--NNCVMEY 405

QY 397 EESNNSLSAEFKHLTLRQRCNGCRANDASLIYTEELHLITFETEVYHQS--TKIDL 454

DB 406 HQA-TQTLAHEFRNMSLK--RIKRSRKAES--VTEKFTLLFDSQPSVGNELVFOV 459

QY 455 EFTSLPWWYISNIGQPNMASILMYNMLTNNKNVNFPTKPIGWDAEVLWSQFS 514

DB 460 KTLSPVWYIVHSQNNATATYLMNARA-EGKRP-FAVPRKVLMPQCEALNNKRYA 517

QY 515 --TTKRGSLIEDLTTLAEKILGPNV---YSCQITMAKFKENNAGKGFSTWMLDNI 568

DB 518 EVQSNNGLTKEINLVFLAQKLFNLSNHLDEDYNSVMSQSFQENENLPGHNYTFWQFDGV 577

QY 569 IDLVKRYIILALWNEGYIMGYSKERERALTSTPPTFTLLRSESSKGGVFTWYEXI 628

DB 578 MEYLAHKLHPHMDGAILGFMKQAHDLINKPDGTFILRSD-SEIGQITLAW--KFD 634

QY 629 SGKTOISVEPYTKQOLNNKSPAEIIMGYKIMDATNIIYSPLYLYPDIPKEAFQKY-- 686

DB 635 SQGRMFWNLMPFTTRDFSIKSLAD-----RLGD-----LNYLIYVPPDKDEVYSKYVT 684

QY 687 ---CRPES-----QHPPE-----ADPGSAAPYLRKTKFCVPTTCSTNTIDLP 725

DB 685 PVCEPATAKKADGYVVKPOIKQVYPEFANASTDAGSGATYMDQ---APSPVCPQA-IYN 740

QY 726 MSPRTLDSLMOFGNCGEAPASAGQFPESI/TFPMDLITSECA 766

DB 741 MYPPNDSVLD--TDGD-----FDLEDITMDVA 765

RESULT 13

ST5B_RAT

AC ST5B_RAT STANDARD: PRT: 786 AA.

DT 01-OCT-1996 (rel. 34, Created)

DT 01-OCT-1996 (rel. 34, Last sequence update)

DT 20-AUG-2001 (rel. 40, Last annotation update)

DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5B.

GN STAT5B.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=96102059; PubMed=8530402;

RA Ripberger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,

RA Fey G.H.;

RT *transcription factors Stat3 and Stat5b are present in rat liver

RT nuclei late in an acute phase response and bind interleukin-6

RT response elements *;

RL J. Biol. Chem. 270:29998-30006(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=NOBLE; TISSUE=Lymph node;

RA Luo G., Yu-Lee L.;

RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.

CC -i- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND

CC ACTIVATION OF TRANSCRIPTION. BINDS TO THE GAS ELEMENT AND

CC ACTIVATES PRL-INDUCED TRANSCRIPTION (BY SIMILARITY).

CC -i- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY

CC MEMBER (BY SIMILARITY).

CC -i- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN

CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).

CC -i- P.TM: TYROSINE PHOSPHORYLATED (BY SIMILARITY).

CC -i- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.

CC -i- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC -----

EMBL: U47686; AAC50491.1; -
 DR HSSP: P42224; 1BF5.
 DR MIM: 604260; -
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; STAT; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR Transcription regulation: DNA-binding; Nuclear protein;
 KM Phosphorylation; SH2 domain.
 FT DOMAIN 589 686
 FT MOD_RES 699 699 SH2.
 FT PHOSPHORYLATION (BY JAKS) (BY
 FT SIMILARITY).
 FT DOMAIN 232 321
 FT CONFLICT 230 230 REQUIRED FOR INTERACTION WITH NMI.
 FT CONFLICT 628 628 S -> P (IN REF. 2).
 FT CONFLICT 717 717 S -> T (IN REF. 2).
 FT CONFLICT 720 720 D -> DA (IN REF. 2).
 FT CONFLICT 720 720 R -> G (IN REF. 2).
 SQ SEQUENCE 786 AA: 89880 MW: E4265C4C8EB824B6 CRC64;

Query Match 21.4%; Score 862.5; DB 1; Length 786;
 Best Local Similarity 30.2%; Pred. No. 1,3e-46;

Matches 248: Conservative 151; Mismatches 311; Indels 111; Gaps 31;

OY 1 MAQNOLOQDTRYLKQHLQYSDTFPMELRQFLAPWIESQDWAYA---ASKESHATLV 56
 DB 1 MAVVIAOQLOQEGALHOMALYQGHFIEVRHYLSQWIESQAMDSYDLDPENIKATQL 60
 OY 57 FNLIGETDOQYSRFLOESNVLYQHNLRIRIKQFLQSRYLEKEMELARIARCLMESRL 116
 DB 61 LEGVQELQKKAEHOGVGEQGLFKLGHATQLOTFYDRCMEVLVRCIRHLILYNQRLV 120
 OY 117 QTAATAOQGGQANHTAAYVTEKQMLEQHLQDYRKRVQDLEQKKVVENLQDDDFNY 176
 DB 121 REANNSSPAGS---LADAMSQKHLQINQTFEELRLVYQDENELKRLQOQYETIIOY 176
 OY 177 -KTLKQ--GDMQDNG---NNQSVTRQKKQOLEQML---TALDQRRSIVSELAQL 224
 DB 177 QESLRIQAFQGLQSLQFQERLSQERLQALQKQVSLQWLEQREQTLQOQRYVELAEKHQKT 236
 OY 225 LSAMEYVQKTLDEELADMKRRPEIACIGGPPNICLDRLNMTLSAEQLOTRQOIKL 284
 DB 237 LQLLKQOITIIIDDLQMKRQQLAGNGPREGSLVQSCXELALITIMONRQOIRRA 296
 OY 285 BELQOKVSYGDPVQVQHRPMLEERIVLEFRNLKMSAFVVEROPCMRPMHREPLVITGVO 344
 DB 297 EHLQCOUHPFG-PVEEMLAEVNATITDIISALVTSTPIIEKQP-----POVLKQTRK 347
 OY 345 FTTKRLVLPPELNYQ---KIKYCIDKQSDVAALGSKRFN-----TLCTNTRKVMNM 396
 DB 348 FAATYRLLVVG-GKLNVHNNPQVQKATIISEQAKSLKNENTRNYSGEIL-NNCCVMEY 405
 OY 397 EESNNGSLSAEKKLTLREORCGNGRANCASLIIVTEELHLITPETEYHOG--LKIDL 454
 DB 406 HQA-IGTISAHRRNNSLK--RIKRSDRGAGES---VTEBEKTIIFESQPSVGNELVROY 459
 OY 455 ETHSLPVPVYINICOMPAMASILMYNMLTNNPKVNFETKPIGTWQVAVLSMWF-S 514
 DB 460 KTLSPVAVVIVHGSQDNNATATVLDNAFA-EPGRVP-FAVPDKVLMPQLCEALMKKFA 517
 OY 515 -TTRRGISTEQLTLAKELIGPG---VNSGCOITWAKCKCKEMAKGGSFWWLNDI 568
 DB 518 EVQSGRGLTKENLVFLAQLKLNSSSHLEDVSGLSVSNQRENRELPGRNVTWQMPFGV 577
 OY 569 IDLVKKYLLALMNGYINGFTSKEREKRALSTKPGFTLRFSSSKGCVFTWVEDDI 628
 DB 578 MEVLAKHLKLRPMNDGAILGIVNKKQAHDLINKPGFTFLRFSD-SELGGSIAM--KFD 634
 OY 629 SGKQIQQVVEYTKQQLNNMSFAELIMGYKIMDATNIVSPVLYLPDIPKEEAFGKYCR 688
 DB 635 SQERFMVNLMPFTTRDSIRSLAD-----RLGD-----LNVLIYVFPDRKDEVYSKYTT 684

OY 689 P--ES-----QEHPE-----ADPGSAAPYLTKPICVPTTCSTNIDLP 725
 DB 685 PVPCEATAKAVDGYVKKPQIKQVYPEFVNASADGCSATYMQD---APSPAVCPQA-HYN 740
 OY 726 MSFRTLDLSLMQFONNCEGAPSAAGQFESTLTPMDLTSQCA 766
 DB 741 MYPQNDPSYLD--TDGD-----FDLEDTMDVA 765
 RESULT 15
 ST5A_HUMAN
 ID ST5A_HUMAN STANDARD: PRT: 794 AA.
 AC P42229;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 5A.
 GN STAT5A OR STAT5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95236292; PubMed=7719937;
 RA Hou J., Schindler U., Henzel W.J., Wong S.C., McKnight S.L.;
 RT *Identification and purification of human Stat proteins activated in
 RT response to interleukin-2.*;
 RL Immunity 2:321-329(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lin J.X., Nietz J., Modi W.S., John S., Leonard W.J.;
 RL Submitted, (EC-19c5) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: CARRIES OUT A DUAL FUNCTION: SIGNAL TRANSDUCTION AND
 CC ACTIVATION OF TRANSCRIPTION. BINDS TO THE GAS ELEMENT AND
 CC ACTIVATES PRL-INDUCED TRANSCRIPTION.
 CC -I- SUBUNIT: FORMS A HOMODIMER OR A HETERODIMER WITH A RELATED FAMILY
 CC MEMBER (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN
 CC RESPONSE TO PHOSPHORYLATION.
 CC -I- PIM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-2, IL-3, IL-7, IL-
 CC 15, GM-CSF, GROWTH HORMONE, PROLACTIN, ERYTHROPOIETIN AND
 CC THROMBOPOIETIN. TYROSINE PHOSPHORYLATION IS REQUIRED FOR DNA-
 CC BINDING ACTIVITY AND DIMERIZATION. SERINE PHOSPHORYLATION IS ALSO
 CC REQUIRED FOR MAXIMAL TRANSCRIPTIONAL ACTIVITY (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L41142; AAA73962.1; -
 DR EMBL: U43185; AAB0589.1; -
 DR HSSP: P42224; 1BF5.
 DR MIM: 601511; -
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001217; STAT.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF01017; STAT; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 KM Transcription regulation; DNA-binding; Nuclear protein;
 KM Phosphorylation; SH2 domain.
 FT DOMAIN 589 686
 FT MOD_RES 694 694 SH2.
 FT MOD_RES 780 780 PHOSPHORYLATION (BY JAKS).
 FT CONFLICT 88 88 PHOSPHORYLATION (BY SIMILARITY).
 FT CONFLICT 88 88 G -> R (IN REF. 2).

